

# Data Architecture



Designed by FreePeek

## Online Education • Certification • Enterprise Solutions

- Data Architecture Fundamentals
- Deciphering Data Architectures
- Architecture and Design for Data Interoperability
- Designing & Implementing Analytics Data Architecture
- Data Modeling Fundamentals
- Operational Data Architecture Part 1: The Operational Data Landscape
- Operational Data Architecture Part 2: Architectural Data Management
- Knowledge Graph Architecture for the Enterprise
- Streaming Data: Concepts, Applications, and Use Cases
- MDM Fundamentals: Architecture and Implementation
- Data Integration Techniques for Designing an ODS
- Introduction to Graph Databases

# Table of Contents

OVERVIEW	2
CIRICULA AT-A-GLANCE	3
CERTIFICATION PROGRAM	4
ENTERPRISE SOLUTIONS	5
COURSE DESTRIPTION	7
OUR INSTRUCTORS	31
ABOUT ELEARNINGCURVE	33
OUR CUSTOMERS	34
CONTACT US	35

---

Data architecture is the art and science of planning and designing the things that are needed to manage data – to collect, organize, store, integrate, quality assure, protect, provide, and use data. Data architecture defines and describes the components, structure, and relationships of an organization’s data assets and data management practices.

The need for adaptable data management architecture has never been more pressing. Yet getting there seems to be more challenging than ever. The frameworks and principles that define and describe how data is organized and managed are rapidly changing. Contemporary data architecture patterns include Modern Data Warehouse, Data Lake, Data Lakehouse, Data Fabric, Data Mesh, and more. Each of these architectural models has advantages and challenges in meeting data management requirements of business operations, analytics, and AI applications. Data architects need knowledge and skills to blend features of multiple architectural approaches into a hybrid architecture, to evolve architecture to meet new requirements, and to support legacy data use cases while building for future capabilities.

Our Data Architecture curriculum includes 11 online courses from world leading experts: Dean Allemang, Angelo Bobak, William McNight, Marc Peco, Kevin Petrie, James Serra, John Singer, and David Wells. Our robust Certified Information Management Professional (CIMP) program builds upon education to certify knowledge and understanding of data architecture. Finally, eLearningCurve’s Enterprise Program is a flexible, scalable, cost-effective solution for teams and enterprises.



# Curricula at-a-Glance

## Data Architecture Fundamentals

**Instructors: Mark Peco & Dave Wells**

This 5-hour course looks at the concepts, principles, and products of data architecture through six different lenses – business alignment, data lifecycle management, data usage, content & structure, processing & storage, and technology.

## Deciphering Data Architectures

**Instructor: James Serra**

This 4-hour course teaches which data architecture frameworks are best suited for specific business needs and how to implement them effectively to ensure data integrity, availability, and usability.

## Designing and Implementing Analytics Data Architecture

**Instructors: Dave Wells & Jed Summerton**

This 4.5-hour course explores modern data management challenges, describes practices & data architecture design patterns, and describes a step-by-step process to get from business requirements to a modern data management architecture that is sustainable and adaptable to the future changes that are sure to come.

## Data Modeling Fundamentals

**Instructor: Dave Wells**

This 4 hour and 45 min. online training teaches practical data modeling skills ranging from traditional relational modeling to key-value, document, graph, and semantic data modeling.

## Operational Data Architecture Part 1: The Operational Data Landscape

**Instructor: Angelo Bobak**

This 5.5-hour course focuses on understanding today's operational data landscape – an essential first step toward modernizing operational data architecture and operational data management.

## Operational Data Architecture Part 2: Architectural Data Management

**Instructor: Angelo Bobak**

This 4-hour course focuses on understanding today's operational data landscape – an essential first step toward modernizing operational data architecture and operational data management.

## Knowledge Graph Architecture for the Enterprise

**Instructor: Dean Allemang**

This 4.5-hour online training course gives businesses the perspective they need to migrate from viewing their enterprise information as a collection of data to viewing it as an ongoing graph of meaningful knowledge.

## Streaming Data: Concepts, Applications, and Technologies

**Instructors: Dave Wells & Kevin Petrie**

This 3-hour course covers the concepts, applications, and business and technical drivers for streaming data adoption, including an in-depth discussion of Apache Kafka.

## MDM Fundamentals: Architecture and Implementation

**Instructor: William McNight**

This 4.5-hour course provides a comprehensive view of the critical elements for MDM success including business, architectural, people, process, project, and technology considerations. Learn from the real-world experiences of the instructor what it takes to make MDM work.

## Data Integration Techniques for Designing an ODS

**Instructor: Angelo Bobak**

This 3-hour online course presents a simple yet thorough process that describes the challenges of building an Operational Data Store (ODS) and the solutions to these challenges.

## Introduction to Graph Databases

**Instructor: John Singer**

This 4-hour online course will provide an overview of property graph database technology and teach the student how to translate business requirements to a property graph database design that can be implemented on any modern property graph database.

## Architecture and Design for Data Interoperability

**Instructor: David Wells**

Data interoperability is quickly becoming a practical alternative to copy-based data integration. Interoperability eliminates the proliferation of data copies and reduces the number and complexity of ETL processes and data pipelines. Interoperability also supports agility and adaptability.

Full course descriptions begin on page 7.

# Certification Programs



## CIMP: Demonstrate Mastery. Achieve Success.

Certification is an important tool for job seekers and for employers seeking to hire the most qualified people. eLearningCurve offers a robust certification program, Certified Information Management Professional (CIMP) that builds upon education to certify knowledge and understanding of information management.

The CIMP — Data Architecture designation makes a clear statement that you have learned from the industry leaders and have demonstrated thorough understanding of data quality by passing several challenging exams.

For the true experts and standard bearers in the industry we offer the second level of CIMP certification - CIMP Ex. To earn the CIMP Ex designation you must demonstrate a combination of great Expertise, Experience, and Excellence.

### WHAT SETS CIMP APART?

**Rigorous exam system:** We go beyond the basics. Rather than testing for knowledge that any industry professional should know, CIMP exams test an in-depth knowledge, comprehensive understanding, and ability to apply various concepts to a problem. You can be proud of your achievement of the CIMP designation, and hiring managers can be sure they are getting a highly knowledgeable employee.

**Education to support certification:** We believe that the best way to ensure success is to combine meaningful industry experience with thorough academic study. To that end, CIMP exams are aligned with our courses, developed and taught by top industry educators and professionals.

**Designed with busy, working professionals in mind:** No time-consuming or costly travel is required to complete coursework or to take your CIMP examinations. All courses and exams are available online. All that's required of candidates is an internet connection and the desire to demonstrate mastery of data governance topics and achieve success.

### HOW DO I ENROLL?

The most convenient and cost-efficient method to enroll in the CIMP program is with one of our Education Packages (see page 5 for details). Each package includes all courses and exams necessary to earn CIMP or CIMP Ex. Alternatively, you can enroll in courses one at a time.



# Enterprise Solutions



**Today more than ever companies are watching expenses and looking for ways to streamline processes, make training convenient, and create a consistent, scalable learning environment.**

Today more than ever companies are watching expenses and looking for ways to streamline processes, make training convenient, and create a consistent learning environment.

eLearningCurve Enterprise is a flexible, convenient, and cost-effective way to train your employees and ensure that all team members have access to information management training they need when they need it. Whether your team or department work in the same office, or are on the other side of the world from each other, you can train them on time and on budget with eLearningCurve Enterprise.

## **Why eLearningCurve Enterprise?**

- Comprehensive educational solution from a single provider
- Employees can take the courses they need when they need them
- Ensure all team members are trained to the same high standard
- Train employees no matter what their geographic location
- Employ a fully scalable education solution
- Minimize disruption to the business
- Maximize your employee training ROI
- Achieve 100% information comprehension
- Get "live" time with our instructors
- Stretch your training budget
- Get solutions for your specific needs

## **When you become an eLearningCurve Enterprise Customer**

We'll work with you to develop educational programs for different roles, positions, teams, departments, and manage and track enrollment of all students in online classes and CIMP exams. We'll track and report educational progress for each student and work with you to meet any specific educational needs.

# eLearningCurve Enterprise Benefits

## **PARTNERSHIP:** *Comprehensive educational solution from a single provider.*

We'll be your educational "partner-for-life" providing employees with continuous information management education they need over the course of their careers.

## **FLEXIBILITY:** *Employees can take the courses they need when they need them.*

Our flexible program allows employees to take the courses they need when they need them to best suit their role, projects, backgrounds or interests.

## **CONSISTENCY:** *Ensure all team members are trained to the same high standard.*

Train your existing team, and set up courses for new hires and transfers. Consider CIMP exams to verify that your employees utilize the same methodology, techniques, and terminology.

## **SCALABILITY:** *Select an Education Partner who truly understands scalability.*

Roll out to a few employees, or your entire organization. Our solution can quickly and effortlessly accommodate groups of all sizes, even if they are geographically dispersed.

## **BREADTH:** *Acquire comprehensive education and certification.*

We offer a full information management education. We have you covered with a comprehensive set of courses, exams, and certifications designed to impart knowledge, test understanding, and validate learning.

## **LOCATION:** *Train employees no matter what their geographic location.*

Overcome geographical barriers to training. You can train your entire team whether they are in the same office, or on the opposite sides of the world. Everyone can access our online courses from any place at any time.

## **LOGISTICS:** *Minimize disruption to the business.*

Our online format allows employees to study from their office or home, allocate full training days, or study an hour a day during lunch breaks.

## **ROI:** *Maximize your employee training ROI.*

No need to worry about paying for flights, hotels and other travel expenses. 100% of what you spend goes towards learning, thus achieving top quality education at a fraction of the cost of in-person training.

## **RESULTS:** *Achieve 100% information comprehension.*

Learn from top industry experts in information management topics. Study at your own pace, listen to the material many times, and test your knowledge through CIMP certification exams.

## **SAVINGS:** *Stretch your training budget.*

We offer various pricing options including volume discounts, pay-as-you-go model with increasing discounts, and other alternatives.

## **"LIVE" INTERACTION:** *Spend time with our instructors.*

Organize question and answer meetings (via Webinar) with course instructors for groups of students who complete online courses.

## **INFORMATION MANAGEMENT 101 Mini-classes.**

As a benefit to our enterprise customers we offer a certain number of complimentary licenses for our 101 mini-classes.

## **CUSTOM COURSEWARE:** *Get solutions for your specific needs.*

Tell us which courses your organization needs the most. We'll work with top instructors in the industry to meet your needs in the most expedient manner.

# Course Descriptions

## Data Architecture Fundamentals

**Instructors: Marc Peco & David Wells**

**Duration: 5 hours**

Data architecture frames how data is managed from the point of creation or collection, through processing of many kinds, to distribution, usage, and business impact. It provides concepts, structures, guidelines, and standards needed for consistency, reliability, resilience, adaptability, and sustainability of data management processes and practices.

This 5-hour online course looks at the concepts, principles, and products of data architecture through six different lenses – business alignment, data lifecycle management, data usage, content & structure, processing & storage, and technology. Putting the pieces together, the course concludes with a look at six steps to architecture design, six methods of architecture implementation, and the importance of continuously evolving data architecture.

### You will learn:

- Multiple, complementary definitions of data architecture
- Why data architecture is needed – both business and technical cases
- How data architecture relates to enterprise architecture, data culture, and data governance
- Six perspectives of data architecture and the underlying concepts of each
- What is produced by architectural activities and processes
- How data architecture is developed, managed, and implemented
- Best practices for data architecture

### This course is geared towards:

- Practicing and aspiring data architects
- CDOs, CIOs, other executives with a role in defining data strategy
- Enterprise, business, systems, technology, analytics, and other architects who work with data architects
- Data engineers, application designers and developers, data systems designers and developers, and others who apply data architecture
- Anyone who needs to collaborate with data architects, and everyone with an interest in data architecture

# Course Outline

## About the Course (5 min)

### Introduction to Data Architecture (49 min)

- Architecture Concepts
- Data Architecture Concepts
- Data Architecture Context
- Data Architecture Perspectives
- Data Architecture Products
- Applied Data Architecture

### Data Architecture Business Perspective (29 min)

- Data Stakeholders
- Data & Business Processes
- Data Domains
- Data Services & Data Products
- Data Monetization

### Data Architecture Lifecycle Perspective (32 min)

- Lifecycle Concepts
- Introducing Data Assets and Data Products
- Operational Data Lifecycle
- Data Assets Lifecycle: Schema on Write & Schema on Read
- Data Products Lifecycle
- Data Architecture and Data Lifecycles

### Data Architecture Usage Perspective (26 min)

- Data Use Cases
- Blending Humans with Technology
- Types of Use Cases
- Role for Data Architecture
- Application Areas and Data Architecture
- Questions Data Architecture Needs to Answer

## Data Architecture Content and Structure Perspectives (37 min)

- Architecture Impact of Data Structure
- Digital Representation of the World
- Building Blocks of Information – A Metaphor
- Event Data and Reference Data
- Measures and Metrics
- Transformed and Refined Data
- Multi-Media Data
- Relational Data Model (Entity Relationship Model)
- Dimensional Data Model
- Non-Relational Data
- NoSQL Data Model

## Data Architecture Processing and Storage Perspective (46 min)

- Operational Systems & Databases
- Data Integration Systems & Databases
- Data Resources
- Data Management Architectures

## Data Architecture Technology Perspective (32 min)

- Establishing a Framework
- Architecture Decisions
- Summary of the Capability Framework
- Role of Technology Architecture

## Putting the Pieces Together (37 min)

- Developing Data Management Architecture
- Implementing Data Management Architecture
- Apply Data Architecture
- Data Architecture Best Practices



# Deciphering Data Architecture

**Instructor: James Serra**

**Duration: 4 hours**

Data fabric, data lakehouse, and data mesh have emerged as viable alternatives to the modern data warehouse. These more recent architecture frameworks have solid benefits, but they're also surrounded by a lot of hyperbole and confusion. This course provides a guided tour of each architecture to help data professionals understand the pros and cons.

The course begins with a review of common data architecture concepts and the traditional data warehouse including its evolution into the modern data warehouse. Then we'll take a detailed look at the rise of data fabrics for seamless data integration and access. Next, we will cover the data lakehouse architecture that combines the best features of data lakes and data warehouses. Lastly, we will study the data mesh, a decentralized data architecture that treats data as a product.

We will delve into the fundamentals of each of these architectures, discussing their structure, use-cases, benefits, and challenges. We will explore how they fit into different business scenarios, their suitability for diverse data types, and the strategies used for their implementation and management. By the end of the course, students will have a clear understanding of which data architecture frameworks are best suited for specific business needs and how to implement them effectively to ensure data integrity, availability, and usability.

## You will learn:

- Concepts that provide working understanding of several data architectures
- The pros and cons of each approach
- How to distinguish data architecture theory from reality
- How to pick the best architecture for your use case
- The differences between data warehouses and data lakes
- Common data architecture concepts to help you build better solutions
- The historical evolution and characteristics of data architectures

## This course is geared towards:

- Practicing and aspiring data architects
- CDOs, CIOs, and other executives with a role in defining data strategy
- Enterprise, analytics, and technology architects who work with data architects
- Data engineers, application designers and developers, data system designers and developers, and others who apply data architecture
- Anyone who needs to collaborate with data architects, and everyone with an interest in data architecture

## Course Outline

### About the Course (2 min)

### Big Data and Data Architectures (23 min)

- Big Data
- Data Architectures

### Common Data Architecture Concepts (85 min)

- Relational Data Warehouse (RDW)
- Data Lake
- Data Storage Solutions
- Approaches to Design
- Approaches to Data Modeling
- Approaches to Data Ingestion

### Modern Data Warehouse (24 min)

- MDM Architecture
- Stepping Stones to the MDW

### Data Fabric (14 min)

- What is Data Fabric? – Part 1
- What is Data Fabric? – Part 2
- Data Fabric High-level Architecture
- Why Data Fabric?
- Data Fabric Drawbacks
- Data Fabric Architecture
- Intelligent Data Fabric

### Data Lakehouse (20 min)

- Data Lakehouse Historical Timeline
- Data Lake: Part 1
- Data Lake: Part 2
- Data Lakehouse High-level Architecture
- Use Cases for Data Lakehouse
- Data Lakehouse Architecture
- Data Lakehouse Needs a Relational Serving Layer
- Concerns Skipping Relational Data Warehouse: Part 1
- Concerns Skipping Relational Data Warehouse: Part 2

### Data Mesh Foundation (41 min)

- Data Mesh Foundation
- Data Mesh Foundation cont.

### Should You Adopt a Data Mesh (25 min)

- Data Mesh Myths
- Data Mesh Myths: Example
- Data Mesh Concerns
- Common Data Mesh Exceptions: Part 1
- Common Data Mesh Exceptions: Part 2
- Should You Adopt a Data Mesh?
- Keys for a Successful Data Mesh
- Data Mesh Future: Part 1
- Data Mesh Future: Part 2

# Designing and Implementing Analytics Data Architecture

**Instructors: Dave Wells & Jed Summerton**

**Duration: 4.5 hours**

Today's business environment requires the use of data and analytics to compete effectively. With continuous advances in data and analytics technologies and capabilities, organizations often struggle to keep up with the changes and to manage data for maximum value and impact. Nearly every organization today is facing the need to rethink and refresh data architecture, yet most continue to work with turn-of-the-century architecture from the BI era. Patching new components onto the surface of legacy architecture—a band aid and duct tape approach—is not sustainable and does a poor job of supporting modern analytics use cases.

Still, many avoid stepping up to modern data architecture because it is complex and difficult. The challenge is two-fold: to clearly define needed business and data capabilities, and to determine how best to weave new capabilities into existing data management practices. This course explores modern data management challenges, describes modern practices and data architecture design patterns, and describes a step-by-step process to get from business requirements to a modern data management architecture that is sustainable and adaptable to the future changes that are sure to come.

## You will learn:

- The reasons that legacy data architectures need to be modernized
- The multitude of requirements for effective analytics data management
- The similarities and differences of Data Lake, Data Fabric, and Data Mesh architectures
- Techniques to identify analytics business capabilities and requirements
- Techniques to identify analytics data capabilities and requirements
- How to apply architectural design patterns and frameworks
- How to adapt reference architectures
- The path from requirements to a well-designed architecture
- Six techniques for architecture implementation

## This course is geared towards:

- Practicing and aspiring data architects
- CDOs, CIOs, and other executives responsible to provide data management leadership
- Enterprise, analytics, and technology architects who work with data architects
- Data engineers and systems engineers
- Designers and developers of data management and analytics systems
- Anyone who needs to collaborate with data architects

## Course Outline

### About the Course (7 min)

#### Legacy Data Management Constraints (15 min)

- Data Sources
- Data Flow and Processing Part 1 & 2
- Data Structure
- Growth and Infrastructure
- Use Cases
- Roles
- Relationships
- Dependencies

#### Analytics and Data Architecture Capabilities (27 min)

- Data and Processing
- Data Platforms
- Data Governance
- Data Consumption
- Modernizing Data Architecture

#### More Requirements for Analytics and Data Architecture (21 min)

- Twelve Requirements
- Thirty Questions

#### Data Architecture Design Patterns (46 min)

- Data Pipeline Patterns
- Data Warehouse and Data Lake Patterns
- Master Data Management Patterns

#### Data Management Architecture Frameworks (43 min)

- Data Lake
- Data Fabric
- Data Mesh
- Data Network
- Data Services

### Defining Your Architecture (44 min)

- Step-by-Step Process
- Business Capabilities and Requirements
- Data Capabilities and Requirements

### Designing Your Architecture (36 min)

- Apply Framework and Design Patterns
- Adapting a Reference Architecture

### Testing and Implementing Your Architecture (23 min)

- Testing the Architecture
- Guiding Principles
- Implementing the Architecture

# Data Modeling Fundamentals

**Instructor: Dave Wells**

**Duration: 4 hours 45 min**

Data modeling is a set of techniques that are fundamental to the processes of understanding, designing, implementing, and curating data. Despite the many declarations that “data modeling is dead” it continues to be an essential part of effective data management. Data Modeling is not dead, but the practices of data modeling are different today than in the past.

Once the domain of database designers and developers, data modeling is now an important skill for data engineers, data scientists, data analysts, application developers, and data curators. Modeling for traditionally structured data is now extended to encompass the variety of big data and NoSQL data types. Semantic data modeling simplifies data integration and is essential to achieve data interoperability. In today’s data modeling practices, long-standing modeling techniques are combined with new and different techniques to adapt to variety in data and data management use cases.

This 4 hours 45 minutes course teaches practical data modeling skills ranging from traditional relational modeling to key-value, document, graph, and semantic data modeling.

## You will learn:

- What data modeling is and why it is important for modern data management
- Data modeling use cases and the roles of data modeling in data management
- Top-down data modeling for data requirements analysis and database design
- Data model reverse engineering to explore, understand, and describe existing data
- Entity-Relationship data modeling techniques
- Multi-dimensional modeling techniques
- NoSQL data modeling techniques including key-value, document, and graph data modeling
- Semantic data modeling techniques including modeling of ontologies and taxonomies

## This course is geared towards:

- Data engineers responsible to design, build, and support databases of all types
- Data engineers responsible to design, build, and support data pipelines
- Data analysts, data scientists, and data engineers who need to investigate, understand, and document data
- Data architects responsible for data standards, data interoperability, and data integration
- Data warehouse and data lake architects, designers, developers, and implementers
- Master data management (MDM) architects, designers, developers, and implementers
- Application systems architects, designers, developers, and implementers
- Anyone with responsibility for or interest in data modeling



# Course Outline

## About the Course (2 min)

### Introductions to Data Modeling (37 min)

- What is Data Modeling?
- Why Data Modeling is Needed?
- Levels of Data Modeling
- Kinds of Data Models
- Module Summary

### Entity-Relationship Modeling (59 min)

- Entity-Relationship Modeling Basis
- Conceptual Modeling
- Logical Modeling
- Physical Modeling
- Module Summary

### Multi-Dimensional Data Modeling (51 min)

- Multi-Dimensional Modeling Basics
- Conceptual Modeling
- Logical Modeling
- Physical Modeling
- Dimension Design Techniques
- Module Summary

### NoSQL Data Modeling (69 min)

- NoSQL Modeling Basics
- Key-Value Data Modeling
- Document Store Data Modeling
- Graph Data Modeling
- Module Summary

### Semantic Data Modeling (64 min)

- Semantic Modeling Basics
- Modeling Ontology
- Modeling Taxonomies
- The Enterprise Semantic Model
- Module Summary

# Operational Data Architecture Part 1: The Operational Data Landscape

**Instructor: Angelo Bobak**

**Duration: 5.5 hours**

Over the past 20 to 30 years the practices of architectural data management have focused primarily on analytical data – architecture for data warehousing and data lakes. Throughout those years, Operational Data Architecture has been largely neglected while the scope, variety, and complexity of operational systems and operational data have expanded substantial.

The operational data landscape – once composed predominantly of transactional systems – now includes automation and IoT systems, as well as, transactional applications. Transactional systems of the past were primarily developed internally with some attention to consistency and connectedness among systems. Today's operational data landscape includes legacy, ERP, SaaS, and custom applications. Operational data platforms encompass mainframe, on-premises servers, cloud, multi-cloud, and mobile devices.

This 5-hour, 22-minute course focuses on understanding today's operational data landscape – an essential first step toward modernizing operational data architecture and operational data management.

## You will learn:

- History and evolution of operational systems
- The variety of operational systems and the roles of each
- The similarities and differences of Data Lake, Data Fabric, and Data Mesh architectures
- The variety of operational data platforms and the characteristics of each
- The variety of data created and managed by operational systems
- Implications of global data variations and mobile data variations
- The roles of MDM and RDM in operational data management

## This course is geared towards:

- Current and aspiring data architects
- Data scientists and data engineers
- Data warehousing and data lake designers and developers
- Data and information systems program and project managers

# Course Outline

## About the Course (2 min)

### Traditional Operational Systems (50 min)

- Transactional Systems
- Automation Systems

### IoT and Data Platforms (54 min)

- Commercial and Industrial IoT
- Operational Data Platforms

### Global Data Architecture Conventions (55 min)

- Data Structures
- Big Data
- JSON Documents
- Sensor & Telemetry Data
- Naming & Data Standards
- Format Standards

### Global Data Architecture Challenges (62 min)

- Global Enterprise Conflict
- Distributed Data
- Duplicate Data
- Conflicting Standards
- Conflicting Data Models
- Data Islands & Gaps
- Homogenous Data vs. Heterogeneous Data
- How to Integrate Heterogeneous and Homogenous Data

### Master Data & Reference Data Examples (57 min)

- The Role of Reference & Master Data
- Master Data
- Application of Master Data
- Reference Data

### Master & Reference Data Challenges (39 min)

- Management, Operational & Technology Challenges
- Data Quality Architecture
- Implementing a Data Quality Technology Strategy
- Implementing a Data Quality Governance Model

## Summary & Conclusions (3 min)

# Operational Data Architecture Part 2: Architectural Data Management

**Instructor: Angelo Bobak**

**Duration: 4 hours**

For several decades, data architecture has focused almost exclusively on the management of analytical data. Most data architecture teams have given little attention to operational data while the scope, variety, and complexity of operational systems and operational data has expanded substantially.

As the operational data landscape has evolved and expanded, data architecture processes and practices have lagged behind. Data sprawl and data disparity increase the cost and complexity of operational data management. These data management challenges inhibit data integration, system interoperability, and business agility. Rethinking operational data architecture is an essential step to overcoming these barriers.

This 4-hour, 14-minute course describes concepts and considerations for rethinking and modernizing operational data architecture. It focuses on positive and proactive data management practices for operational data, and the underlying architectural processes and patterns to enable and reinforce those practices.

**Prerequisite:** Operational Data Architecture, Part 1: The Operational Data Landscape

## You will learn:

- Architectural concepts, constructs, and techniques to manage data sprawl and data disparity
- Concepts and management practices for distributed data, homogeneous data, and heterogeneous data
- Concepts and management practices for data conflicts including semantic and schema conflicts
- Master data management (MDM) and reference data management (RDM) principles and practices
- Operational data architecture patterns including operational data store (ODS), data hubs, data brokers, and more
- Design techniques for adaptable and sustainable architecture

## This course is geared towards:

- Current and aspiring data architects
- Data scientists and data engineers
- Data warehousing and data lake designers and developers
- Data and information systems program and project managers

# Course Outline

## About the Course (2 min)

### Distribution, Data Silos, and Data Conflicts (47 min)

- Managing Distributed Data
- Managing Homogenous & Heterogeneous Data Automation Systems
- Managing Conflicting Database Schema

### Data Integration Architecture (51 min)

- Managing Master & Reference Data
- Managing Semantic Models

### Technical Operational Data Architecture (60 min)

- Operational Data Store
- Publish/Subscribe Paradigm
- Operational Data Hub
- Service Oriented Architecture
- Related Data Integration Technologies

### The Physical Architecture (50 min)

- Case Study
- Architecture Requirements
- Modeling Your Processes

### Implementation & Management (40 min)

- Identifying Architecture Issues
- Current State, Future State, & Gap Analysis
- Implementation Timeline
- Managing the Architecture

### Summary, Conclusions, & Next Steps (4 min)



# Knowledge Graph Architecture for the Enterprise

**Instructor: Dean Allemang**

**Duration: 4 hours 37 min**

Businesses today need to have comprehensive control and understanding of the data both within the enterprise and from outside. Regulatory pressure in many industries requires that an enterprise be able to find all relevant data about customers, products, business relations, and in fact any aspect of the business quickly and accurately.

Advances in data analytics have shown that the more meaningful data an enterprise can bring to bear, the more accurately business drivers can be predicted. For this reason, many businesses are migrating from viewing their enterprise information as a collection of data to viewing it as an ongoing graph of meaningful knowledge.

## You will learn:

- What a Knowledge Graph is, what are its uses, and how does it provide business value
- Different approaches to Knowledge Graphs, and when to use them
- Knowledge Graph standards and their importance in governance and application sustainability
- Categories of tools for building, deploying, and maintaining knowledge graphs, and how they fit together into a knowledge graph deployment
- The relationship between knowledge and data in a knowledge graph, and how to manage them both
- Best Practices for data publication, use and re-use

## This course is geared towards:

- Data integration architects, designers, and developers
- Data and technology architects
- Chief Data Officers

**Course exam is included in the course price.**

## Course Outline

### About the Course (7 min)

#### Definitions and Business Case (55 min)

- History of Knowledge Graphs
- Knowledge Graphs in Various Industries - Finance
- Knowledge Graphs in Various Industries - Media
- Knowledge Graphs in Various Industries - Agriculture
- Knowledge Graphs in Various Industries - Pharmaceuticals
- Data Management Challenges in the Enterprise

#### Knowledge Graph Supporting Technologies (40 min)

- Features of Graph Data Systems
- Approaches to Graphs as Data
- Property Graph Data Capabilities
- Graph Data Queries
- Graph Data Visualization
- Graph Data Systems
- Standardizing Data

#### Semantic Technology Fundamentals (58 min)

- Knowledge Graph Stack
- RDF Brings Data Together
- RDFS Enhance Data with Types & Properties
- OWL Provides Precision Logic to Describe Models
- SPARQL Let's Ask Questions About the Data
- Example Knowledge Graph Capabilities Using Semantic Web Standards
- Using SKOS for Knowledge Management
- The Meaning of Meaning

### Knowledge Graph Enterprise Framework (35 min)

- Application Vs. Enterprise Data
- The Enterprise Data Jungle
- Sustainable Extensibility
- Enterprise Data Community
- Prerequisites for Distributed Data
- Common Reference
- Connecting References
- Semantic Alignment
- Things vs. Strings
- Tools and Components
- Applications of Knowledge Graphs

### Modeling Methodology and Architecture (33 min)

- Modeling Methodology and Architecture
- Terminology Sources
- Guidelines for Terminology

### Implementation Fundamentals of Knowledge Graph (47 min)

- Example Implementation
- Incremental Development
- Knowledge Graph Implementation
- Presentation Approach
- Select a Use case
- Inventory Data Sources
- Draw on Reference Ontologies
- Identify Controlled Vocabulary
- Map Metadata to Ontologies
- Materialize Data as Needed
- Build Queries to Respond to Business Questions
- Repeat from Step 3 with New Data
- How Can This Fail?
- Summary: Incremental Development

# Streaming Data: Concepts, Applications, and Technologies

**Instructors: Dave Wells & Kevin Petrie**

**Duration: 3 hours**

Streaming data is data that continuously flows from sources such as IoT devices, sensors, GPS devices, server and security logs, and clickstreams from mobile apps and websites—typically high-volume data moving at high speed. The analytics opportunities with IoT and application data streams are abundant, but the value of streaming technology is not limited to native data streams. In today's fast paced business world, the need for fast data is pervasive and tacit acceptance of high-latency data is rapidly diminishing. Streaming as an alternative to batch ETL is a practical way to meet the demand for fast data.

Change Data Capture (CDC) is a category of technology that captures data about changes made to a database – inserts, updates, and deletes – and makes that data available to downstream processing such as data pipelines that flow to data warehouses and data lakes. CDC can be combined with streaming to accelerate data flow and reduce data latency.

Apache Kafka is a widely adopted open source technology for stream processing. It is an open source, distributed streaming platform that is used to move high volumes of data in real time. Building data pipelines with Kafka requires knowledge of Kafka architecture, components, and processes. You'll need to know the actions and responsibilities of data producers and of data consumers, as well as the capabilities for cluster management, data connections, and APIs. Integrating Kafka or other streaming technologies into your data ecosystem is an important consideration.

## You will learn to:

- The business and technical drivers for streaming data adoption
- Data pipeline processing patterns and the advanced patterns that are possible with streaming
- How microservices and edge computing work with IoT data streams
- Use case patterns and a variety of use cases for streaming data
- Five kinds of Change Data Capture (CDC) and the strengths and weaknesses of each
- The concept and applications of streaming first architecture
- Kafka architecture and essential components
- Kafka data and process flow
- The roles and functions of Kafka broker, data producers, and data consumers
- Cluster management, data connections, and APIs with Kafka
- Integrating streaming into the data ecosystem

## This course is geared towards:

- Data and analytics leaders and managers
- Data and analytics architects
- Data scientists
- Data engineers
- Data governance professionals who need to understand the opportunities and implications of streaming data
- Anyone with a desire to know how streaming is changing the data management landscape

## Course Outline

### About the Course (5 min)

#### The Need for Speed (21 min)

- Batch Vs. Real Time
- The Speed of Business
- The Speed of Data Parts 1-3
- Fast Data Drives the Consumer World
- Fast Data Drives the Business World
- Fast Data Drives the Analytic World
- The Business Case for Fast Data

#### Moving Data through Pipelines (61 min)

- Data Pipeline Processing
- Data Stream Processing
- IoT and Edge Computing

#### Change Data Capture (10 min)

- Data Management - A New Streaming Use Case
- What is CDC?
- Date Stamps for Change Detection
- Diff Comparisons for Change Detection
- Database Triggers for Change Detection
- In-Database Change Data Capture
- Log Based Change Data Capture
- Log Based CDC for Data Warehousing and Data Lakes

#### Streaming First Architecture for Data Management (17 min)

- Modern Enterprise Data Requirements
- Essential Characteristics of Streaming Technology
- Stream Processing Technologies
- Kafka Stream Processing Basics
- Replacing ETL with CDC and Streaming
- Streaming First Architecture
- Stream to Data Warehousing
- Stream to Slowly Changing Dimensions
- Stream to Data Lake
- Stream to Analytic Applications
- Moving to Streaming First Architecture

### Data Streaming with Apache Kafka (11 min)

- Introduction to Apache Kafka
- Kafka Predecessors

#### Kafka Architecture (19 min)

- Apache Kafka Architecture
- Essential Components Part 1-2
- Data and Process Flow Part 1-2
- Kafka Record - Key Components
- Kafka Broker Overview
- Cluster Management with Apache

### Data Stream Ecosystems and Use Cases (40 min)

- Data Streaming Use Cases
- Streaming Integration with Data Ecosystems
- Kafka Alternatives from Cloud Service Providers
- Amazon Kinesis
- Azure Event Hub
- Google Cloud PUB/SUB

# MDM Fundamentals: Architecture and Implementation

**Instructor: William McNight**

**Duration: 4.5 hours**

Proliferation of heterogeneous systems creates a pressing need for data sharing and data consistency. When many different systems collect data about master entities – customers, products, suppliers, employees, accounts, etc. – you can be certain that you'll find inconsistencies, conflicts, and confusion. At best, conflict and confusion leads to waste and inefficiency in business process. More severe consequences include damaged credibility and reputation when errors and inconsistencies are visible to customers, suppliers, and employees. Today's complex business and information systems must synchronize master data. That is the role and purpose of Master Data Management (MDM) systems.

MDM is not a casual endeavor. It is a complex data management challenge that requires a formal and well-managed program. The unique challenges of an MDM program are often not apparent even to seasoned data management professionals. The complexities of managing identities and resolving conflicts among disparate data sources make MDM an ambitious undertaking that must address business, architectural, people, process, project, and technology dimensions to succeed.

This course provides a comprehensive look at the elements of an MDM program and the key success factors for MDM.

## You will learn:

- The what and why of Master Data Management (MDM)
- A variety of architectural approaches to MDM and how to determine which is the best fit for your MDM program
- The human dimension of MDM including roles and responsibilities of sponsors, managers, analysts, architects, designers, and developers
- The state of MDM technologies along with techniques and guidelines for tool selection
- The process dimension of MDM including impacts upon business processes and information management processes
- The project perspective of MDM including organizing and executing the activities of planning, requirements analysis, design, development, testing, data migration, and implementation.

## This course is geared towards:

- MDM Program and Project Managers
- MDM Analysts, Designers, and Developers
- Business Data Owners, Data Stewards, and Data Consumers
- Data Architects
- Information Systems Project Managers
- Data Integration Program and Project Managers
- Data Stewards, Data Governance Professionals, and Data Quality Practitioners



## Course Outline

### About the Course (7 min)

○

### Introduction (54 min)

- MDM Overview
- MDM Justification and Outcomes
- Master Data Management for Customers

### MDM Architecture (90 min)

- Architecture Approaches
- Conforming Dimensions for the Enterprise
- Data Quality
- Syndicated Data
- Additional Considerations

### MDM Tool Selection (48 min)

- Process Considerations
- Proof of Concept and Final Selection

### MDM Project Execution (77 min)

- MDM Project Management
- MDM Project Roles and Responsibilities
- Organizing and Planning for MDM Success
- Case Study

# Data Integration Techniques for Designing an ODS

**Instructor: Angelo Bobak**

**Duration: 4 hours**

For several decades, data architecture has focused almost exclusively on the management of analytical data. Most data architecture teams have given little attention to operational data while the scope, variety, and complexity of operational systems and operational data has expanded substantially.

As the operational data landscape has evolved and expanded, data architecture processes and practices have lagged behind. Data sprawl and data disparity increase the cost and complexity of operational data management. These data management challenges inhibit data integration, system interoperability, and business agility. Rethinking operational data architecture is an essential step to overcoming these barriers.

This 4-hour, 14-minute course describes concepts and considerations for rethinking and modernizing operational data architecture. It focuses on positive and proactive data management practices for operational data, and the underlying architectural processes and patterns to enable and reinforce those practices.

**Prerequisite:** Operational Data Architecture, Part 1: The Operational Data Landscape

## You will learn:

- Architectural concepts, constructs, and techniques to manage data sprawl and data disparity
- Concepts and management practices for distributed data, homogeneous data, and heterogeneous data
- Concepts and management practices for data conflicts including semantic and schema conflicts
- Master data management (MDM) and reference data management (RDM) principles and practices
- Operational data architecture patterns including operational data store (ODS), data hubs, data brokers, and more
- Design techniques for adaptable and sustainable architecture

## This course is geared towards:

- Current and aspiring data architects
- Data scientists and data engineers
- Data warehousing and data lake designers and developers
- Data and information systems program and project managers

## Course Outline

### About the Course (2 min)

○

### Distribution, Data Silos, and Data Conflicts (47 min)

- Managing Distributed Data
- Managing Homogenous & Heterogeneous Data Automation Systems
- Managing Conflicting Database Schema

### Data Integration Architecture (51 min)

- Managing Master & Reference Data
- Managing Semantic Models

### Technical Operational Data Architecture (60 min)

- Operational Data Store
- Publish/Subscribe Paradigm
- Operational Data Hub
- Service Oriented Architecture
- Related Data Integration Technologies

### The Physical Architecture (50 min)

- Case Study
- Architecture Requirements
- Modeling Your Processes

### Implementation & Management (40 min)

- Identifying Architecture Issues
- Current State, Future State, & Gap Analysis
- Implementation Timeline
- Managing the Architecture

### Summary, Conclusions, & Next Steps (4 min)

# Introduction to Graph Databases

**Instructor: John Singer**

**Duration: 4 hours**

Entity Relationship modeling and relational databases have dominated the IT scene since the '80s, becoming the de facto standard approach for data persistence. However, the ubiquitous relational database has waned with the advent of NoSQL and big data technologies. Today's data architect must master a new database technology - graph database - that has emerged with a solid set of use cases based on mathematical graph theory and graph algorithms.

This online course will provide an overview of property graph database technology and teach the student how to translate business requirements to a property graph database design that can be implemented on any modern property graph database.

## You will learn:

- To understand property graphs as a new database model
- To compare and contrast property graphs with other database models
- To understand how a graph database fits into the overall data ecosystem
- How to properly design a property graph data model
- Property Graph database modeling and design best practices for entity definition, relationship definition, and modeling complex objects

## This course is geared towards:

- Data Architects that need to understand how graph database fits into the overall persistence architecture
- DBA's and Data modelers expanding into graph database
- Data Science and Data analysts interacting with graph databases
- Big Data Managers and decision makers

## Course Outline

### About the Course (7 min)

○

### Introduction to the Graph Data Model (28 min)

- Graph Theory
- Anatomy of a Node
- Anatomy of a Relationship
- Properties and Paths

### The Graph Database Eco-System (65 min)

- Graph Database
- Graph Analytics
- Semantic Web Graph Database

### Agile Graph Data Modeling (39 min)

- Graph Data Modeling Overview
- Agile Graph Data Modeling
- Graph Data Modeling Process
- Working with Users

### Entity Modeling (42 min)

- Representing Things
- Describing Things
- Categorizing Things Part 1 & 2
- Entity Definition Best/Worst Practices

### Relationship Modeling (38 min)

- Representing Connections
- Naming Relationships
- Relationship Direction
- Describing Relationships
- Relationship Best/Worst Practices

### Complex Object Modeling (40 min)

- Complex Object Modeling
- Resolving Hypergraphs Part 1 & 2
- Complex Objects
- Data Structures-Linked Lists and Trees
- Managing Slowly Changing Dimensions



# Architecture and Design for Data Interoperability

Instructors: David Wells

Duration: 3 hours

Data integration has been the primary method of resolving data disparity for decades. However, integration brings challenges with multiple copies of data and many data pipelines. Data interoperability is quickly becoming a practical alternative to copy-based data integration. Interoperability eliminates the proliferation of data copies and reduces the number and complexity of ETL processes and data pipelines. Interoperability also supports agility and adaptability. Data is not force-fitted into a rigid schema but instead positioned as independent components that can be configured to fit many different use cases.

Making data interoperable begins with data architecture to identify interoperability patterns and weave them into your existing data management architecture. Knowing what patterns you'll use - APIs, data products, data contracts, semantic mapping, open links, etc. - is an essential first step. Next, you'll need to identify the technologies needed to enable each pattern and fit them into your technology stack. With those pieces in place, you're ready to provide design guidelines and begin migration planning.

## You will learn:

- What data interoperability is and why it is important
- The full scope of data management architecture - operational and analytical data
- How operational data management practices contributes to data friction and technical debt
- How analytical data management practices become barriers to data interoperability
- How data interoperability mitigates data friction and reduces technical debt
- How existing data management architecture can be adapted for interoperability
- Semantic data modeling and mapping as the foundation of data interoperability
- Interoperability patterns using APIs, data products, data contracts, and data virtualization
- How a semantic layer works as the nucleus of data interoperability

## This course is geared towards:

- Practicing and aspiring data architects of all types - enterprise data architects, analytical data architects, operational data architects, and application data architects
- Practicing and aspiring data engineers of all types - data product engineers, data pipeline engineers, and database engineers
- Data strategists including CDOs, CIOs, and other executives with roles in defining data strategy
- Data warehouse and data lake architects, designers, developers, and implementers
- Master data management (MDM) architects, designers, developers, and implementers
- Application systems architects, designers, developers, and implementers

## Course Outline

### About the Course (3 min)

○

### Operational Data Architecture (37 min)

- The Operational Data Landscape
- Managing Operational Data
- An Architectural View of Data Management
- Module Summary

### Data Interoperability (43 min)

- The *What* and *Why* of Data Interoperability
- Architectural Barriers to Data Interoperability
- Module Summary

### Data Semantic and Data Interoperability (41 min)

- Data Semantics & Data Management
- Semantic Data Modeling
- Semantic Data Mapping
- Module Summary

### The Enterprise Semantic Layer (56 min)

- Semantic Layer Perspectives
- APIs
- Data Products
- Data Contracts
- Data Virtualization
- Interoperability in Data Management Architecture
- Technology Overview
- Module Summary

# Our Instructors

## Dean Allemang

Dean Allemang has over 20 years of experience with research and deployment of semantic solutions, including expert systems, knowledge management systems, and rule-based systems. Throughout most of his career, he has been interested in how knowledge can be formalized and used to improve the ways in which people carry out their work.

## Angelo Bobak

Angelo Bobak is a seasoned data architecture professional and published author with over 20 years of experience in business intelligence, data architecture, data modeling, master data management, and data quality. Currently he is working at ATOS Origin NA as a Director/Senior Data Architect in the areas of global master data management, data integration, and data quality.

## William McKnight

William McKnight is president of McKnight Consulting Group, which includes the service lines of master data management, IT assessment, data warehousing, and business intelligence. He functions as strategist, lead enterprise information architect, and program manager for sites worldwide. William is a former Information Technology Vice President for a Fortune 50 company, a former engineer of DB2 at IBM, and holds an MBA from Santa Clara University.

## Mark Peco

Mark is an experienced consultant, educator, practitioner, and manager in the fields of BI and process improvement. He provides vision and leadership to projects, operating and creating solutions at the intersection of business and technology. Mark is involved with clients working in the areas of strategy development, process improvement, data management, and business intelligence.

## Kevin Petrie

Kevin Petrie's passion is deciphering what technology means to practitioners. He has invested more than 20 years in technology as a strategy analyst, instructor, product marketer, and services leader. A frequent public speaker and accomplished writer, Kevin has nearly a decade of experience in data management and analytics. Kevin is currently Senior Director of Product Marketing at Attunity, a provider of data integration software based in Burlington, MA, where he also runs sales and new hire training.

## James Serra

James Serra works at Microsoft as a big data and data warehousing solution architect. He is a thought leader in the use and application of Big Data and advanced analytics, including data architectures. He is a prior SQL Server MVP with over 35 years of IT experience, a popular blogger ([JamesSerra.com](http://JamesSerra.com)) and speaker, having presented at major events, including SQLBits, PASS Summit, Data Summit, and the Enterprise Data World Conference. James is the author of the book “Deciphering Data Architectures: Choosing Between a Modern Data Warehouse, Data Fabric, Data Lakehouse, and Data Mesh”.

## John Singer

John Singer has 4 decades of experience in a variety of data architecture-related roles. John’s accomplishments include implementing metadata management solutions, data modeling processes and governance, master data management solutions, and an ITIL-based CMDB, combining architecture, business, and IT metadata in a comprehensive solution. John is currently focusing on property graph technologies and is the founder and CEO of NodeEra Software.

## David Wells

Dave Wells is an advisory consultant, educator, and author dedicated to building meaningful connections throughout the path from data to business value. He works at the intersection of information management and business management, driving business impact through data literacy, data culture, data architecture, data analytics, and active data management. More than fifty years of data and business management experience provide unique perspectives about the connections among business, information, data, and technology.

Dave is both an instructor and the Education Director for eLearningCurve, an advisory consultant and contributing author at Eckerson Group, and an instructor and speaker with Dataversity.



Designed by FreePeek

## About eLearningCurve

eLearningCurve offers comprehensive online education programs in various disciplines of information management. With eLearningCurve, you can take the courses you need when you need them from any place at any time. Study at your own pace, listen to the material many times, and test your knowledge through online exams to ensure maximum information comprehension and retention.

eLearningCurve also offers two robust certification programs: CIMP & CDS. Certified Information Management Professional (CIMP) builds upon education to certify knowledge and understanding of information management. Certified Data Steward (CDS) is a role-based certification designed for the fast growing data stewardship profession.

Finally, eLearningCurve's Enterprise Program is a flexible, scalable, cost-effective solution for teams and enterprises.

# Our Customers

eLearningCurve has students in almost every country in the world, including many enterprise customers.



- Variety of customers from small project teams to large enterprises
- Enterprise customers typically Fortune 500 and Global 1000 companies
- All major industries are represented

## WHAT OUR CUSTOMERS ARE SAYING...



**The courses are well laid out, build on each other, and are rich in practical content and advice.**

-- Steve Lutter, CIMP Data Quality, DM and Metadata, IM Foundations, Business Intelligence, Data Governance, MDM, United States



**It is evident that a thorough and considerable effort has gone into the preparation of this program.**

-- Alfredo Parga O'Sullivan, CIMP Ex Data Quality,



**The ability to take the courses at my own pace and at a time suitable for me was of great help.**

-- Geeta Jegamathi, CIMP Data Quality, India

# Contact Us

## **Director, Enterprise Solutions**

Arkady Maydanchik  
[Arkadym@elearningcurve.com](mailto:Arkadym@elearningcurve.com)

## **Director, Course Design and Production**

Michelle Johnson  
[Mjohnson@elearningcurve.com](mailto:Mjohnson@elearningcurve.com)

## **Director, Education**

Dave Wells  
[David.Wells@elearningcurve.com](mailto:David.Wells@elearningcurve.com)

## **Director, Technology**

Varya Belyaevskaya  
[Admin@elearningcurve.com](mailto:Admin@elearningcurve.com)

## **Director, Marketing**

Aleksandra Labuda  
[Alabuda@elearningcurve.com](mailto:Alabuda@elearningcurve.com)

## **Customer Support**

[Support@elearningcurve.com](mailto:Support@elearningcurve.com)

**Phone:** 1-630-242-1659

### **Mailing Address:**

305 Midwest Club Parkway  
Oak Brook, IL 60523



# Resellers

## **DENMARK**

### **Right Training**

#### **Contact:**

Jan Sørensen  
+45 31 45 04 99  
[jms@righttraining.dk](mailto:jms@righttraining.dk)

## **SOUTH AFRICA & SUB-SAHARAN AFRICA**

### **Master Data**

#### **Contact:**

Gary Allemann  
+27 11 485 4856  
[gary@masterdata.co.za](mailto:gary@masterdata.co.za)