

Metadata Management



Online Education • Certification • Enterprise Solutions

- **Metadata Management Fundamentals**
- **Curating and Cataloging Data**
- **Data Profiling**
- **Data Quality Assessment**
- **Data Modeling Fundamentals**
- **Knowledge Graph Architecture for the Enterprise**
- **Introduction to Graph Databases**
- **Data Quality Scorecard**
- **Data Parsing, Matching, and De-duplication**
- **DW and BI Data Modeling**
- **The Data Model Scorecard**
- **Data Architecture Fundamentals**
- **Architecture and Design for Data Interoperability**

Table of Contents

OVERVIEW	2
CURRICULA AT-A-GLANCE	3
CERTIFICATION PROGRAMS	5
ENTERPRISE SOLUTIONS	6
COURSE DESCRIPTIONS	8
OUR INSTRUCTORS	21
OUR CUSTOMERS	24
ABOUT ELEARNINGCURVE	25
CONTACT US	26

Data modeling and metadata management expanded rapidly in recent years. Exponential increase in the number and complexity of databases and interfaces between them, as well as huge rise in importance of efficient data governance and data quality management changed the landscape of metadata management. Similarly, emergence of dimensional data shook the foundations of data modeling. Today we see new technologies driving further change in the modeling of structured data. Beyond structured data we find new challenges in unstructured data – text, images, voice, video, and more.

Our Metadata Management curriculum includes online courses from world leading experts: Dean Allemang, Angelo Bobak, Mike Bracket, Steve Hoberman, Kathy Hunter, William McKnight, Arkady Maydanchik, Olga Maydanchik, John Singer, Henrik Sørensen, Rick Sherman, and Dave Wells. Our robust Certified Information Management Professional (CIMP) program builds upon education to certify knowledge and understanding of data quality. Finally, eLearningCurve's Enterprise Program is a flexible, scalable, cost-effective solution for teams and enterprises.

WHAT PEOPLE ARE SAYING ABOUT ELC



I learned so much out of the courses that I wanted to continue my learning even beyond the basic requirement of my current job.

-- Jagmeet Singh, CIMP Ex - Data Modeling & Metadata Management, Data Quality, Data Governance, IM Foundations, MDM USA

Curricula-at-a-Glance

Full course descriptions begin on page 8.

Metadata Management Fundamentals

Instructors: Arkady Maydanchik and Dave Wells

This 4-hour course is designed to provide the foundational metadata knowledge needed by anyone who has data management roles and responsibilities.

Curating and Cataloging Data

Instructor: David Wells

This 3-hour online training course will explore how curating and cataloging work together to meet the data needs of business and data analysts, to provide self-service data to complement self-service analytics, and to realize the promise of democratizing data analytics.

Data Profiling

Instructor: Arkady Maydanchik

Data profiling is the process of analyzing actual data and understanding its true structure and meaning. It is one of the most common and important activities in information management. This 5-hour course teaches all practical skills necessary to succeed in a data profiling initiative.

Data Quality Assessment

Instructor: Arkady Maydanchik

This 5-hour course gives comprehensive treatment to the process and practical challenges of data quality assessment. Special attention is paid to the architecture and functionality of the data quality metadata warehouse.

Data Modeling Fundamentals

Instructor: Dave Wells

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

Knowledge Graph Architecture for the Enterprise

Instructor: Dean Allemang

In this 4-hour 37-min you will learn how to find all relevant data about customers, products, business relations and in fact any aspect of the business quickly and accurately, and how to view the collection of data as ongoing graph of meaningful knowledge.

Introduction to Graph Databases

Instructor: John Singer

This 4-hour online training 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.

Data Quality Scorecard

Instructors: Olga Maydanchik

This 5-hour course gives comprehensive treatment to the processes and practical challenges of data quality scorecarding.

Data Parsing, Matching, and De-duplication

Instructors: Kathy Hunter, William

McKnight, Henrik Sørensen

To take advantage of the worldwide marketplace, businesses need to manage data globally. This reality poses very specialized and unique kinds of problems in data management. In this 3-hour course you will learn to identify and avoid the pitfalls of global information.

BI & DW Data Modeling

Instructor: Rick Sherman

This 4-hour course includes a mix of data modeling concepts, best practices, applications and practical examples that will help you build effective data warehouse and business intelligence applications.

The Data Model Scorecard

Instructor: Steve Hoberman

We often build data models quickly and with the singular goal of database design. This 3-hour course presents Data Model Scorecard®, which provides the tools needed to measure and manage data model quality.

Data Architecture Fundamentals

Instructor: David Wells, Mike Brackett, Arkady Maydanchik

This 4-hour course is designed to provide foundation knowledge about the most commonly used metadata management, data modeling, and data profiling techniques.

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.



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 Metadata Management designation will make a clear statement that you have learned from the industry leaders and have demonstrated thorough understanding of data modeling and metadata 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 metadata management 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. 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

Metadata Management Fundamentals

Instructor: David Wells and Arkady Maydanchik

Duration: 4 hours

Deriving value from data depends extensively on understanding the data and sharing knowledge among everyone who works with data. Sharing data knowledge is the core purpose of metadata. Just as you need financial data to manage financial resources, you need metadata to manage data resources. In today's data-driven world, the importance of managing data is certainly on par with that of managing finances.

This online training course is designed to provide the foundational metadata knowledge needed by anyone who has data management roles and responsibilities. It covers metadata basics such as the types and purposes of metadata, and explores core metadata disciplines of data modeling, data profiling, and data cataloging. Metadata roles in data governance, stewardship, security, quality, and analysis are explained.

You will learn:

- The scope and complexities of metadata management
- The roles of data models as metadata and the basics of data modeling
- The role of data profiling in metadata management and the basics of data profiling methods
- The roles of data catalogs in metadata management and the fundamentals of data curation and data cataloging
- Metadata dependencies of business processes, IT projects, data governance, data quality, business intelligence, self-service data, business analytics, and data science

This course is geared towards:

- Anyone with data management roles and responsibilities
- Data stewards and data governance practitioners and participants
- Data curators and data catalog administrators
- Data and database analysts and designers
- Data quality professionals and practitioners
- Aspiring data modelers who need to start with the basics
- Anyone with a role in information management who needs to understand data or help others to understand data

Course Outline

About the Course (8 min)

Understanding Data (15 min)

- Views of Data
- Projects Flow
- Describing the Data Meaning
- Describing the Data Constraints
- Describing the Data Relationships
- Describing the Data

Metadata Management (57 min)

- Metadata
- Metadata Management Processes
- Using Metadata
- Metadata Tools and Technologies

Data Modeling (24 min)

- Data Modeling Defined
- The Data Modeling Process
- Supplemental Models & Additional E-R Concepts
- Dimensional Data Modeling

Data Profiling (47 min)

- What is Data Profiling?
- Myth and Reality of Data Profiling
- Profiling Techniques
- Profiling Challenges
- Role of Profiling
- People and Technology

Data Curation and Cataloging (31 min)

- Data Curation
- Data Cataloging
- Metadata and the Catalog

Metadata Management for BI and Data Science (49 min)

- The Metadata Muddle
- Data Science and Metadata
- Data Provenance and Data Lineage
- Ontology and Taxonomy

Curating and Cataloging Data

Instructors: Dave Wells

Duration: 3 hours

As the world of data management grows and changes, the roles and participants in data ecosystems must adapt. With the convergence of several influences – big data, self-service analytics, and self-service data preparation – we need to actively manage the inventory of self-service data. Data curation is both a data inventory management process and a data governance activity. The data curator is responsible to oversee a collection of data assets and make it available to and findable by anyone who needs data. Cataloging maintains the collection of metadata that is necessary to support browsing, searching, evaluating, accessing, and securing datasets.

This 3-hour online training course will explore how curating and cataloging work together to meet the data needs of business and data analysts, to provide self-service data to complement self-service analytics, and to realize the promise of democratizing data analytics.

You will learn:

- How to define and identify disparate data.
- The concepts, responsibilities, and skills of data curation
- The role of the data curator in data governance and the differences between a data curator and a data steward
- The needs of data seekers and the ways that curating and cataloging help to meet
- The purpose, content, and uses of a data catalog
- The state of data cataloging tools and technology
- Guidelines for getting started with data curating and cataloging

This course is geared towards:

- Business and IT leaders struggling with the paradoxes of modern data management
- Analytics and BI designers and developers who are dependent on fresh and relevant data for every analytics use case
- Data management professionals at all levels from architects to engineers
- Data governance professionals – especially data stewards who need to adapt to the changing world of modern data management

Course Outline

About the Course (5 min)

Self-Service Data (9 min)

- Governance and Self-Service
- How We Got Here
- Why Self-Service Data?

Data Curation (44 min)

- Data Curation
- Why Data Curation?
- The Data Curator
- Data Lifecycles and Curation
- Curating Big Data
- Getting Started with Data Curation

Data Cataloging (30 min)

- Definitions
- Why Data Cataloging?
- Metadata and the Catalog
- Data Catalog Tools
- Getting Started with Data Cataloging

Evaluating Data Catalog Tools (42 min)

- The Business Case for a Data Catalog
- Kinds of Data Catalogs
- Evaluation Criteria

Beyond Self-Service Data (44 min)

- Information, Technology, and Data
- The Enterprise Data Marketplace
- EDM Architecture
- EDM Data Services
- Data as a Service

Data Profiling

Instructor: Arkady Maydanchik

Duration: 5 hours

Data profiling is the process of analyzing actual data and understanding its true structure and meaning. It is one of the most common and important activities in information management. Data profiling is the first critical step in many major IT initiatives, including implementing a data warehouse, building an MDM hub, populating metadata repository, as well as operational data migration and integration. It is also the key ingredient to successful data quality management.

While proliferation of commercial tools made data profiling accessible for most information management professionals, successful profiling projects remain elusive. This is largely because the tools allow gathering large volumes of information about data, but offer limited means and guidelines for analysis of that information.

In this online training course you will learn all practical skills necessary to succeed in a data profiling initiative.

You will learn:

- The what, why, when, and how of data profiling
- Various data profiling techniques, from simple column profiling to advanced profiling methods for time-dependent and state-dependent data
- How to efficiently gather data profiles
- How to analyze the data profiling information and ask the right questions about your data
- How to organize data profiling results
- How to perform dynamic data profiling and identify changes in data structure and meaning

This course is geared towards:

- Data quality practitioners
- MDM practitioners
- Metadata management practitioners
- IT and business analysts involved in data management
- Those responsible for implementation and maintenance of various data management systems

Course Outline

About the Course (7 min)

Introduction to Data Profiling (44 min)

- What is Data Profiling?
- Myth and Reality of Data Profiling
- Profiling Techniques
- Profiling Challenges
- Role of Profiling
- People and Technology

Column Profiling (89 min)

- Introduction
- Basic Counts
- Value Frequency Charts
- Value Distribution Characteristics
- Value Distribution

Profiling Time-Dependent Data (58 min)

- Introduction
- Timeline Profiling
- Timestamp Pattern Profiling
- Multi-Dimensional Profiling
- Event Dependency Profiling

Profiling State-Transition Models (49 min)

- Introduction
- Data Structures for State-Dependent Data
- Profiling Techniques

Other Profiling Techniques (65 min)

- Subject Profiling
- Relational Integrity Profiling
- Attribute Dependency Profiling
- Dynamic Data Profiling

Data Quality Assessment

Instructors: Arkady Maydanchik

Duration: 6 hours

More and more companies initiate data quality programs and form data stewardship groups every year. The starting point for any such program must be data quality assessment. Yet in absence of a comprehensive methodology, measuring data quality remains an elusive concept. It proves to be easier to produce hundreds or thousands of data error reports than to make any sense of them.

This online training course gives comprehensive treatment to the process and practical challenges of data quality assessment. It starts with systematic treatment of various data quality rules and proceeds to the results analysis and building aggregated data quality scorecard. Special attention is paid to the architecture and functionality of the data quality metadata warehouse.

You will learn:

- The what, why, when, and how of data quality assessment
- How to identify and use data quality rules for assessment
- How to ensure completeness of data quality assessment
- How to construct and use a data quality scorecard
- How to collect, manage, maintain, warehouse and use data quality metadata

This course is geared towards:

- Data quality practitioners
- Data stewards
- IT and business analysts and everyone else involved in data quality management

Course Outline

About the Course (9 min)

Introduction (53 min)

- Why Assess Data Quality
- Business Value of Data Quality Assessment
- Types of Data Errors
- Data Quality Assessment Approaches
- How Rule-Driven Approach Works
- Project Planning
- Project Steps

Data Quality Rules Overview (63 min)

- Attribute Domain Constraints
- Relational Integrity Constraints
- Complex Data Relationships

Rules for Historical Data (56 min)

- Historical Data Overview
- Timeline Constraints
- Value Pattern Rules
- Rules for Event Histories
- Rules for State-Dependent Objects

Finding Data Errors (76 min)

- Discovering Data Quality Rules
- Implementing Data Quality Rules
- Building Rule Catalog
- Building Error Catalog
- Fine-Tuning Data Quality Rules

Aggregate Data Quality Scores (66 min)

- School Report Card Example
- A First Look at DQ Scorecard
- Defining Aggregate Scores
- Score Tabulation

Building Data Quality Scorecard (61 min)

- Basic Scorecard Example
- Recurrent Data Quality Assessment
- Database and Enterprise-Wide DQ Scorecard

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.

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

Knowledge Graph Architecture for the Enterprise

Instructor: Dean Allemang

Duration: 4 hours 37 minutes

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 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 Lets You 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

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:

- The fundamental concepts and practices of Business Intelligence
- The fundamental concepts and practices of Business Analytics
- Concepts and principles of Predictive Analytics
- Concepts and principles of Data Mining
- Concepts and principles of Data Science
- The roles of data analytics in a data-driven organization

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)

Module 1. 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

Data Quality Scorecard

Instructor: Olga Maydanchik

Duration: 5 hours

Data quality scorecards have become very popular and many organizations are starting to build them. What they have found is that the path to a meaningful and useful DQ Scorecard is riddled with traps and obstacles.

This online training course gives comprehensive treatment to the processes and practical challenges of data quality scorecarding.

It starts with a few real, live use cases that showcase what a scorecard can do for a company when done right. Systematic treatment of various DQ scorecard challenges is given. Then the course proceeds to the ins and outs of the successful DQ scorecard, from the underlying data model to the effective processes that need to be set up in order to produce the scorecard efficiently. Multiple examples to illustrate every important point are provided in the class.

You will learn:

- The methodology behind data quality metrics calculations
- The best way to organize data quality related metadata collected during typical data quality projects
- Effective data visualization techniques to depict data quality measurements
- Typical pitfalls that accompany data quality scorecard implementation and how to avoid them
- How to achieve scorecard adoption and usage by the business users

This course is geared towards:

- Data quality practitioners
- Data stewards and data governance practitioners
- IT analysts, business analysts, and everyone else involved in data quality management
- Developers tasked with DQ Scorecard creation

Course Outline

About the Course (3 min)

Case Studies (46 min)

- What is Data Quality Assessment?
- What is a Data Quality Scorecard?
- Data Quality Scorecard Case Study 1: Improving the Efficiency of the Risk-Weighted Asset (RWA) Calculation Process (Financial Company)
- Data Quality Scorecard Case Study 2: Data Quality Impact on Catastrophe Risk Modeling (Insurance Company)

Data Quality Score Calculation Methods (37 min)

- Averages Method For Score Cards
- Record Level Score Calculations
- Subject Level Score Calculations
- Score Types Comparisons
- Score Decomposition By Business Dimensions
- Business Dimensions Versus Subjects

Data Modeling Considerations Part 1 (55 min)

- Why DQMDW?
- DQMDW Components
- Case Study DQMDW For Property Insurance Company
- DQMDW: Critical Data Elements Catalog
- DQMDW: Rule Catalog

Data Modeling Considerations Part 2 (68 min)

- DQMDW: Subject Master and Business Dimensions Master
- DQMDW: Error Catalog
- Error Details - Storage Options
- Rule Error Output - Advanced Examples
- DQMDW: Score Catalog
- DataMarts For DQ Visualization

Building A Data Quality Scorecard Process (61 min)

- Process Overview
- Step 1 Define The DQ Assessment and DQ Scorecard Scope
- Step 2: Populate The Staging Area
- Step 3: Prepare The Data
- Step 3A: Add Record ID Step
- Step 3B: Fill Dataset AsOfDate
- Step 3C: Create/Update Subject Master List
- Step 3D: Create/Update Business Dimension Master Lists
- Step 4: Perform Data Profiling
- Step 5: Create and Code Data Quality Rules
- Step 6: Run DQ Rules
- Step 7: Move Rule Execution Results into DQMDW
- Step 8: Calculate Aggregate Scores
- Step 9: Examine DQ Scores and DQ Rule Results
- Step 10: Fine-Tune Data Quality Rules

Data Quality Scorecard Demo (35 mins)

- Overall Score Analysis
- Summary

Data Parsing, Matching and De-duplication

Instructors: Kathy Hunter, William McKnight, Henrik Sørensen

Duration: 4 hours, 20 minutes

Data parsing, standardization, matching, and de-duplication are the cornerstones of successful Master Data Management (MDM). They are also critical parts of successful data quality programs, and are key steps in building data warehouses as well as any data integration and consolidation initiatives. You could say that today few organizations can function effectively without implementing data parsing and matching processes often in many data domains.

This need is further magnified if your company has gone global and plans to create databases that combine name- and address-related data from all corners of the world. Managing global information effectively takes specialist knowledge and the ability to show consideration for the differences that exist throughout the world. Worldwide there are more than 10,000 languages, 130 address formats, 36 personal and hundreds of business name formats. All of these variables are further complicated by the need to respect national and regional cultures. Failure to consider formats, styles, and cultures has huge impact on quality of data and quality of business relationships.

This online training course is aimed at data quality and master data management (MDM) professionals as well as those responsible to work with global information. The field is broad and the details are many. The purpose of this course is to provide a broad and in-depth review of data parsing, standardization, matching, and de-duplication techniques, as well as extensive overview of specific problems and solutions when dealing with global data.

You will learn:

- Data parsing, standardization, matching, and de-duplication techniques
- How to find and use external reference data
- How data parsing and matching contribute to improving data quality, MDM, and data warehousing
- Which data domains, entities and data elements may benefit from data parsing and matching
- Challenges of global data and ways to overcome these challenges

This course is geared towards:

- Master data management professionals
- Data quality professionals
- Information architects
- Developers of data warehousing systems
- Business professionals who work with global data

Course Outline

About the Course (12 min)

Introduction (17 min)

Implementation Fundamentals (70 min)

- Parsing and Standardization
- Introduction to Data Matching
- Data Matching Techniques
- Data Matching Destinations
- Evaluating Data Matching Tools

External Reference Data (45 min)

- External Data Sources
- Syndicated Customer Data
- Syndicated Product Data
- Using the Web
-

Challenges of Global Data (58 min)

- Introduction to Global Information
- Global Data: What You Need to Know
- Variations by Country and Region
- Cultural and Legal Impacts
- Characters and Diacritics

Overcoming the Challenges of Global Data (59 min)

- Data Profiling
- Consistent Data Structures
- Preparing Global Data for Effective Use

DW and BI Data Modeling

Instructor: Rick Sherman

Duration: 4 hours, 15 minutes

A well designed data model is the cornerstone to building business intelligence and data warehouse applications that provide significant business value.

Effective data modeling results in transforming data into an enterprise information asset that is consistent, comprehensive and current. Data is transformed from operational or source systems into a data warehouse and often data marts or OLAP cubes for analysis. This course provides the fundamental techniques to designing the data warehouse, data marts or cubes that enable business intelligence reporting and analytics.

This online training course discusses the two logical data modeling approaches of Entity-Relationship (ER) and dimensional modeling. ER modeling is used to establish the baseline data model while dimensional modeling is the cornerstone to Business Intelligence (BI) and Data Warehousing (DW) applications. These modeling techniques have expanded and matured as best practices have emerged from years of experience in data modeling in enterprises of all sizes and industries. These techniques improve the business value of the data, enhance project productivity and reduce the time to develop applications. This course includes a mix of concepts, applications and practical examples.

You will learn:

- The basics of Entity-Relationship (ER) and dimensional modeling
- the benefits and applicability of Dimensional Data Modeling
- how to create Dimensional Data Models for BI and DW applications
- how to learn more about Data Modeling

This course is geared towards:

- Beginning Data Modelers
- Business Analysts and Architects
- Database Administrators and Analyst
- IT Managers, Project Managers
- Application Development Project Team Members
- People involved in design and maintenance of Data Warehousing and Business Intelligence applications
- People involved in data quality or data governance processes

Course Outline

About the Course (6 min)

Introduction to Data Modeling (75 min)

- Data Modeling Overview
- Entity-Relationship Modeling Overview
- Normalization

Dimensional Modeling Basics (88 min)

- What is Dimensional Modeling?
- Facts
- Dimensions
- Schemas
- Entity-Relationship vs. Dimensional Modeling
- Purpose of Dimensional Modeling
- Fact Tables
- Dimensional Modeling Vocabulary

Advanced Dimensional Modeling (84 min)

- Hierarchies
- Slowly Changing Dimensions
- Rapidly Changing Dimensions
- Casual Dimensions
- Multi-Valued Dimensions
- Snow flaking
- Junk Dimensions
- Value Brand Reporting
- Heterogeneous Products
- How Swappable Dimensions
- Too Few or Too Many Dimensions
- Benefits of Dimensional Modeling

The Data Model Scorecard

Instructors: Steve Hoberman

Duration: 3 hours

A frequently overlooked aspect of data quality management is that of data model quality. We often build data models quickly, in the midst of a development project, and with the singular goal of database design. Yet the implications of those models are far-reaching and long-lasting. They affect the structure of implemented data, the ability to adapt to change, understanding of and communication about data, definition of data quality rules, and much more. In many ways, high-quality data begins with high-quality data models.

This online training course presents Steve Hoberman's Data Model Scorecard®, which provides the tools needed to measure and manage data model quality.

You will learn:

- The importance of having an objective measure of data model quality
- The categories that make up the scorecard including correctness, completeness, structural soundness, flexibility, standards, and model consistency
- How to apply the scorecard to different types of models
- Techniques to strengthen data models, including model reviews, model substitutes (screens, prototypes, sentences, spreadsheets and reports), and the use of automated tools to enforce modeling best practices and standards
- How to introduce the scorecard into a development methodology and your company culture

This course is geared towards:

- Data Modelers
- Data Analysts
- Data Architects
- Data Stewards
- Database Administrators

Course Outline

About the Course (8 min)

Scorecard Need (43 min)

- Why Measure Data Model Quality
- Traditional Review Methods
- Archer vs. Data Modeler
- Enter the Scorecard

Scorecard Categories (67 min)

- Category 1 - Model Type
- Category 2 - Correctness
- Category 3 - Completeness
- Category 4 - Structure
- Category 5 - Abstraction
- Category 6 - Standards
- Category 7 - Readability
- Category 8 - Definitions
- Category 9 - Consistency
- Category 10 - Data

Scorecard in Practice (54 min)

- Introducing the Scorecard into your Organization
- Scorecard Challenges
- Scorecard Tips
- Applying the Scorecard

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

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.

Mike Brackett

Mike Brackett has been in the data management field for over 40 years, during which he has developed many concepts and techniques for designing applications and managing data resources. He is the originator of the common data architecture concept, the data resource framework, the data naming taxonomy, the five-tier five-schema concept, the data rule concept, the BI value chain, the data resource data concept, and the architecture-driven data model concept, as well as new techniques for /integrating disparate data.

Steve Hoberman

Steve Hoberman is a trainer, consultant, and writer in the field of data modeling. He taught his first data modeling class in 1992 and has taught over 10,000 people about data modeling and BI techniques. He has presented at over 50 international conferences, and has been selected to deliver keynote addresses at major industry conferences in North America and in Europe. Steve is a columnist and frequent contributor to industry publications and the author of several data modeling books.

Kathy Hunter

Kathy's interest in data began when she was studying for her Summa Cum Laude Mathematics degree. Later she instituted an Information Quality Division at One2One where her team recovered £10 million in lost earnings. She went on to build a global data management solution set at Harte-Hanks which provided data management capabilities to multinational organizations with data from as many as 238 countries. A popular speaker at information quality events, Kathy is known for her pragmatic approach toward complicated topics.

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.

Arkady Maydanchik

For more than 20 years, Arkady Maydanchik has been a recognized leader and innovator in the fields of data quality and information integration. In 1997, Arkady founded Arkidata Corporation, which was among one of the first companies dedicated solely to data quality management. Since 2004, Arkady has dedicated his efforts toward education and the creation of a mature data quality profession. He is a frequent speaker at various conferences and seminars, author of the best-selling *Data Quality Assessment* book, contributor to many journals and online publications, and a co-founder of eLearningCurve.

Olga Maydanchik

Olga Maydanchik is an experienced practitioner and educator in the field of information management. As a part of the Chief Data Offices in Citi, AIG, Deutsche Bank, and Voya Financial, Olga was focused on designing and implementing the enterprise-wide Data Quality, Master Data Management, Metadata Management, and Analytics programs. Olga is a member of the Enterprise Data Management Council and actively participated in the Data Management Capability Assessment Model and Ontology design work streams.

Rick Sherman

Despite the risk of sounding like an old geezer, Rick Sherman will admit that he's been doing data warehousing since before it was even called data warehousing. Rick is the founder of Athena IT Solutions, a Boston area business intelligence and data warehousing consulting firm that provides solutions for customers of all sizes and types of industries. His hands-on experience includes a wide range of data integration tools. Rick also teaches data warehousing, data integration, and business intelligence for a master's degree program at Northeastern University's graduate school of engineering. He is a published author of more than a hundred articles.

Rick Sherman

Despite the risk of sounding like an old geezer, Rick Sherman will admit that he's been doing data warehousing since before it was even called data warehousing. Rick is the founder of Athena IT Solutions, a Boston area business intelligence and data warehousing consulting firm that provides solutions for customers of all sizes and types of industries. His hands-on experience includes a wide range of data integration tools. Rick also teaches data warehousing, data integration, and business intelligence for a master's degree program at Northeastern University's graduate school of engineering. He is a published author of more than a hundred articles.

John Singer

John Singer has 4 decades of experience in a variety of data architecture-related roles. John's past 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.

Henrik Sørensen

Henrik Liliendahl Sørensen has over 30 years of experience working with master data management and data quality and is a charter member of the International Association of Information and Data Quality. Currently, Henrik works with master data management at Tata Consulting Services and as Practice Manager at Omikron Data Quality, besides writing on a well-trafficked blog about data quality, master data management, and the art of data matching.

David Wells

Dave Wells is a consultant, teacher, and practitioner in the field of information management. He brings to every endeavor a unique and balanced perspective about the relationship between business and technology. This perspective, refined throughout a career of more than thirty-five years that has encompassed both business and technical roles, helps align business and information technology in the most effective ways. Dave is a frequent contributor to trade publications and is a co-author of the book *BI Strategy: How to Create and Document*. He also speaks at a variety of industry events.

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, Ireland



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



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.

Contact Us

Director, Enterprise Solutions

Arkady Maydanchik

Arkadym@elearningcurve.com

Director, Course Design and Production

Michelle Johnson

Mjohnson@elearningcurve.com

Director, Education

Dave Wells

David.Wells@elearningcurve.com

Director, Technology

Varya Belyaevskaya

Admin@elearningcurve.com

Director, Marketing

Aleksandra Labuda

ALabuda@elearningcurve.com

Customer Support

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

SOUTH AFRICA & SUB-SAHARAN AFRICA**Master Data****Contact:**

Gary Allemann

+27 11 485 4856

gary@masterdata.co.za