Master Data Management

Online Education • Certification • Enterprise Solutions

➤ MDM Fundamentals: Architecture and Implementation
➤ Data Parsing, Matching & De-duplication
➤ Data Governance Fundamentals
➤ Metadata Management Fundamentals
➤ Data Integration Fundamentals and Best Practice
➤ Data Quality Assessment
➤ Ensuring Data Quality in Data Integration
➤ Data Quality Scorecard
➤ Data Profiling
➤ Data Virtualization

TABLE OF CONTENTS

Master Data Management (MDM) is a core component of modern data management. MDM is a critical discipline that focuses on uniformity, accuracy, consistency, semantic integrity of shared reference data across the enterprise. A single, consistent view of master entities—customers, products, accounts, employees, etc.—that span multiple business processes is important to ensure both data quality and process integrity. Customer relationship management (CRM) is difficult without customer data integration, and a 360 view of customers is impossible to achieve without actively managing customer data. Similarly, supply chain management is difficult without integrated and consistent product, supplier, and partner data.

These are but two examples of the reasons that MDM is a critical data management capability. As enterprises strive to become data driven and data use cases expand with the adoption of self-service analytics and data science applications, the role and reach of master data expands rapidly. In the age of digital transformation MDM is no longer a "nice to have" capability. It is an essential competency for business success.

WHAT PEOPLE ARE SAYING ABOUT ELC

I found the courses to be very well organized in terms of content and delivery; the exams were challenging and the best thing about this online learning is that you can listen to the content at your own convenient time.

-- Purvi Ramchandani, CIMP Ex – Data Quality, CIMP Ex – MDM, USA

CURRICULA AT-A-GLANCE

Full course descriptions begin on page 7.

MDM Fundamentals: Architecture and Implementation
Instructor: William McKnight
This 4.5 hour course provides a comprehensive look at the elements of an MDM program and the key success factors for MDM.

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 4.5-hour course you will learn to identify and avoid the pitfalls of global information.

Data Governance Fundamentals
Instructors: Maria Villar, Theresa Kushner, Dave Wells
This 5-hour course provides an overview of the disciplines of governing data, covers the essential components of an enterprise-wide program, and outlines a roadmap to execute a successful data governance program.

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. It covers metadata basics such as the types and purposes of metadata, and explores core metadata disciplines of modeling, profiling, and cataloging data.

Data Integration Fundamentals and Best Practice
Instructor: Dave Wells
This online training course discusses architectures, requirements, methods, roles and activities of data integration that can be applied to achieve successful data integration projects for a variety of applications and circumstances.

Data Quality Assessment
Instructor: Arkady Maydanchik
This 6-hour 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.

Ensuring Data Quality in Data Integration
Instructor: Arkady Maydanchik
Without a comprehensive data quality monitoring program bad data spread like viruses through countless real-time and batch interfaces. This 5-hour course discusses various practices that can be put in place to maintain high data quality through data integration.

Data Quality Scorecard
Instructor: Arkady Maydanchik
Data quality scorecards have become very popular and many organizations are starting to build them. This online training course gives comprehensive treatment to the processes and practical challenges of data quality scorecarding.

Data Profiling
Instructor: Olga 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 Virtualization
Instructor: Dave Wells
Data virtualization is a core component of next-generation data integration architectures, techniques, and technology. This 3-hour online training course will introduce you to the concepts, techniques, and capabilities of data virtualization.

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 Master Data Management designation will make a clear statement that you have learned from the industry leaders and have demonstrated thorough understanding of master data management 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 master data 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 rack 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.

SCALABLITY: 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..

CUSTOME 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.

MDM Fundamentals: Architecture and Implementation

Instructor: William McKnight

Duration: 4 hours and 30 minutes

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

MDM Architecture (90 min)
- o Architecture Approaches
- o Conforming Dimensions for the Enterprise
- o Data Quality
- o Syndicated Data
- o Additional Considerations

MDM Tool Selection (48 min)
- o Process Considerations
- o Proof of Concept and Final Selection

MDM Project Execution (77 min)
- o MDM Project Management
- o MDM Project Roles and Responsibilities
- o Organizing and Planning for MDM Success
- o Case Study

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

Overcoming the Challenges of Global Data (59 min)
o Data Profiling
o Consistent Data Structures
o Preparing Global Data for Effective Use

Data Governance Fundamentals

Instructors: Theresa Kushner, Maria C. Villar and Dave Wells
Duration: 5 hours

Data governance is an emerging, cross-functional management program that treats data as an enterprise asset. It includes the collection of policies, standards, processes, people, and technology essential to managing critical data to a set of goals.

Data governance also includes the oversight necessary to ensure compliance and to manage risk. A data governance program can be tailored to match an organization's culture, information maturity, priorities, and sponsorship.

This online training course provides an overview of the disciplines of governing data, covers the essential components of an enterprise-wide program, and outlines a roadmap to execute a successful data governance program. In addition to the extensive overview, the course makes data governance real and tangible by illustrating the concepts, principles, and practices using a case study of data governance in a customer intelligence initiative.

You will learn:
- What data should be governed
- Why data governance is important
- Basic concepts, principles, and practices of a data governance program
- Where and how to start a data governance program
- People and tools that enable a data governance program
- Techniques to measure success of a data governance program
- Governance of big data and cloud applications

This course is geared towards:
- Individuals who implement a data governance program
- Individuals who participate in a data governance program
- Business data stewards
- Information professionals who want to learn about this emerging area

Course Outline

About the Course (7 min)

Introduction to Data Governance (33 min)
- o Data Governance 101
- o Why Govern Data?
- o What Data Should Be Governed?
- o Business Drivers for Data Governance

Implementation Fundamentals (78 min)
- o Selecting Data and Setting Goals
- o Standards, Policies, Processes, People, and Technology
- o Managing and Measuring Data Governance

Case Study (63 min)

Data Governance of Emerging Solutions (65 min)
- Big Data
- Cloud Applications

Modernizing Data Governance (53 min)

- Overview
- The Data Quake: From Stable to Volatile
- New Data Governance Challenges
- Curating and Cataloging Data
- Rethinking Data Governance Practices
- Technologies and Modern Data Governance
- Module Summary

Metadata Management Fundamentals

Instructor: Dave 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)
- o Views of Data
- o Projects Flow
- o Describing the Data Meaning
- o Describing the Data Constraints
- o Describing the Data Relationships
- o Describing the Data

Metadata Management (57 min)
- o Metadata
- o Metadata Management Processes
- o Using Metadata
- o Metadata Tools and Technologies

Data Modeling (24 min)
- o Data Modeling Defined
- o The Data Modeling Process
- o Supplemental Models & Additional E-R Concepts
- o Dimensional Data Modeling

Data Profiling (47 min)
- o What is Data Profiling?
- o Myth and Reality of Data Profiling
- o Profiling Techniques
- o Profiling Challenges
- o Role of Profiling
- o People and Technology

Data Curation and Cataloging (31 min)
- o Data Curation
- o Data Cataloging
- o Metadata and the Catalog

Metadata Management for BI and Data Science (49 min)
- o The Metadata Muddle
- o Data Science and Metadata
- o Data Provenance and Data Lineage
- o Ontology and Taxonomy

Data Integration Fundamentals & Best Practices

Instructor: Dave Wells

Duration: 5 hours

Integrated data is at the heart of many business and technical disciplines today. Data warehousing, operational data integration, and master data management focus on integration as a key part of managing data as an asset. Business intelligence, performance management, and business analytics depend on integrated data to meet business requirements for management and decision-making information. Legacy system replacement, ERP implementation, and application integration all have integrated data dependencies. Integration is important, but it is challenging to understand data sources, select and apply integration techniques, and design and deliver integrated databases.

This online training course discusses architectures, requirements, methods, roles and activities of data integration that can be applied to achieve successful data integration projects for a variety of applications and circumstances.

You will learn:

➢ Fundamental concepts, principles, and terminology of data integration
➢ Common methods of data integration with attention to techniques, timing, and integration process automation
➢ How to perform the essential steps of data integration including requirements definition, data capture, data transformation, and data delivery
➢ Data integration techniques and technologies including ETL, ELT, virtualization, and federation
➢ Techniques for source-to-target mapping and data transformation
➢ Roles, purpose, and variations of data integration architecture including architectural constructs for data warehousing, master data management, and operational data integration
➢ Business and technical roles, responsibilities, knowledge, and skills that are central to data integration projects and processes

This course is geared towards:

➢ Data integration architects, designers, and developers
➢ Business intelligence practitioners, project managers, and architects
➢ Data warehousing practitioners, project managers, and architects
➢ MDM practitioners, project managers, and architects
➢ ERP implementers and system integrators
➢ Business subject experts and data subject experts with roles in data integration projects and processes

Course Outline

About the Course (5min)
Data Integration Concepts (44 min)
  o Data Integration Defined
  o Data Integration Dependent Programs
  o Data Integration Projects

Data Integration Methods (49 min)
  o Data Integration Techniques
  o Data Integration Frequency
  o Data Integration Systems
  o Data Integration Challenges
  o Data Integration Activities

Understanding Data (61 min)
  o Identifying Data Sources
  o Profiling Data
  o Qualifying Data Sources
  o Documenting Data Sources

Integrating Data (78 min)
  o Integration Requirements
  o Data Capture
  o Data Transformation
  o Data Delivery

Data Integration Architecture (37 min)
  o Architecture Concepts
  o Data Warehousing Architecture
  o MDM Architecture
  o Operational Data Integration Architecture

Roles & Responsibilities (28 min)
  o Knowledge and Skills
  o Understanding the Data
  o Getting the Data
  o Transforming the Data
  o Delivering the Data
  o Using the Data

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

Ensuring Data Quality in Data Integration

Instructor: Arkady Maydanchik
Duration: 5 hours

Corporate data universe consists of numerous databases connected by countless real-time and batch data interfaces. The data continuously move about and change. The databases are endlessly redesigned and upgraded, as are the programs responsible for the data integration. The typical result of these dynamics is that information systems get better, while data quality deteriorates. Without a comprehensive data quality monitoring program bad data spread like viruses.

This online training course discusses various practices that can be put in place to mitigate the problem and maintain high data quality through data integration.

You will learn:
- ➢ The data quality challenges that are inherent in data integration
- ➢ The critical role of data quality monitoring in data integration
- ➢ Specific techniques to monitor and manage quality for batch data integration
- ➢ Use of Statistical Process Control (SPC) methods in monitoring data quality
- ➢ The impacts of change on data quality and techniques to address those impacts
- ➢ How an enterprise integration hub can be applied to managing data quality

This course is geared towards:
- ➢ Data integration practitioners
- ➢ Data quality practitioners
- ➢ Data warehousing practitioners
- ➢ MDM practitioners
- ➢ Others in the trenches involved in design, development, and maintenance of data integration systems

Course Outline

About the Course (8 min)

Introduction (60 min)
- o Data Integration Basics
- o Data Quality in Data Federation
- o Data Quality in Data Consolidation
- o Data Quality in Real-Time Interfaces
- o Data Quality in Batch Interfaces
- o Beyond Monitoring

Data Consolidation and Cleansing (60 min)
- o Data Quality in Data Conversion
- o Data Cleansing
- o Data Quality in Data Consolidation

Error Monitors in Batch Interfaces (77 min)
- o Monitoring Techniques
- o Batch Integrity Rules
- o Master Data Integrity Rules
- o Dynamic Integrity Rules
- o Monitor Management
- o Error Correction
- o Root Cause Analysis
- o Statistical Process Control

Change Monitors in Batch Interfaces (95 min)
- o Examples
- o Change Monitor Defined
- o Which Metrics to Monitor
- o Implementing the Monitors
- o Heuristic Monitors
- o Basic Statistical Monitor
- o Handling Non-Stationary Data

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

Data Quality Score Calculation Methods (37 min)
o Averages Method For Score Cards
o Record Level Score Calculations
o Subject Level Score Calculations
o Score Types Comparisons
o Score Decomposition By Business Dimensions
o Business Dimensions Versus Subjects

Data Modeling Considerations Part 1 (55 min)
o Why DQMDW?
o DQMDW Components
o Case Study DQMDW For Property Insurance Company
o DQMDW: Critical Data Elements Catalog
o DQMDW: Rule Catalog

Data Modeling Considerations Part 2 (68 mins)
o DQMDW: Subject Master And Business Dimensions Master
o DQMDW: Error Catalog
o Error Details – Storage Options
o Rule Error Output – Advanced Examples
o DQMDW: Score Catalog
o DataMarts For DQ Visualization

Building A Data Quality Scorecard Process (61 mins)
o Process Overview
o Step 1 Define The DQ Assessment And DQ Scorecard Scope
o Step 2: Populate The Staging Area
o Step 3: Prepare The Data
o Step 3A: Add Record ID Step
o Step 3B: Fill Dataset AsOfDate
o Step 3C: Create/Update Subject Master List
o Step 3D: Create/Update Business Dimension Master Lists
o Step 4: Perform Data Profiling
o Step 5: Create and Code Data Quality Rules
o Step 6: Run DQ Rules
o Step 7: Move Rule Execution Results into DQMDW
o Step 8: Calculate Aggregate Scores
o Step 9: Examine DQ Scores and DQ Rule Results
o Step 10: Fine-Tune Data Quality Rules

Data Quality Scorecard Demo (35 mins)

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:

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

This course is geared towards:

1. Data quality practitioners
2. MDM practitioners
3. Metadata management practitioners
4. IT and business analysts involved in data management
5. Those responsible for implementation and maintenance of various data management systems

Course Outline

About the Course (7 min)

Introduction to Data Profiling (44 min)
1. What is Data Profiling?
2. Myth and Reality of Data Profiling
3. Profiling Techniques
4. Profiling Challenges
5. Role of Profiling
6. People and Technology

Column Profiling (89 min)
1. Introduction
2. Basic Counts
3. Value Frequency Charts
4. Value Distribution Characteristics
5. Value Distribution

Profiling Time-Dependent Data (58 min)
1. Introduction
2. Timeline Profiling
3. Timestamp Pattern Profiling
4. Multi-Dimensional Profiling
5. 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 Virtualization

Instructor: Dave Wells
Duration: 3 hours

The work of data integration has become increasingly complex in recent years. Business needs for real-time and low latency data, expanded uses of unstructured data, and accelerated interest in big data analytics are but a few of the trends that change the data integration landscape. Extract-transform-load (ETL) processing was sufficient for the once relatively simple task of combining data from multiple transactional databases was to build a data warehouse, operational data store, or master data hub. Today's data integration challenges go well beyond the capabilities of ETL technologies with needs to integrate enterprise data with external data, Web data, clickstream data, end-user data, big data, cloud data, and more. To meet these new requirements, data integrators need more tools in the integration toolbox. Data virtualization doesn't replace ETL; it complements ETL and offers new tools to meet new integration needs.

Data virtualization is a core component of next-generation data integration architectures, techniques, and technology. This online training course will introduce you to the concepts, techniques, and capabilities of data virtualization. It will prepare you to expand your data integration capabilities, deliver business-speed information, and make the most of recent advances in data integration technology.

You Will Learn:
➢ Data virtualization definitions, concepts, and terminology
➢ Business case and technical rationale for data virtualization
➢ Foundational principles of virtualization – abstraction, views, and services
➢ How to extend the data warehouse with virtualization
➢ How virtualization is applied for unstructured data, big data, and cloud data challenges
➢ How to mix and match virtualization with ETL technology to optimize data integration architectures and processes

This course is geared towards individuals who:
➢ BI, MDM, and data warehousing program and project managers
➢ Data integration architects, designers, and developers
➢ Data and technology architects

Course Outline

About the Course (6 min)

Data Virtualization Concepts and Principles (29 min)
- o Overview
- o Data Virtualization Basics
- o Why Data Virtualization
- o The Data Virtualization Foundation
- o Review

Data Integration Architecture (19 min)
- o Overview
- o Integration Architecture Concepts
- o Reference Architectures
- o Integration Architecture Examples
- o Review

Data Virtualization in Integration Architecture (49 min)
- o Overview
- o Virtualization in Data Integration Projects
- o Data Virtualization Use Cases
- o Data Warehousing Use Cases
- o Data Federation Use Cases
- o MDM and EIM Use Cases
- o More Data Virtualization Applications
- o Practical Data Virtualization
- o Review

Data Virtualization Platforms (20 min)
- o Overview
- o Platform Requirements
- o Platform Capabilities
- o Platform Variations
- o Some Platform Vendors
- o Review

Implementing Data Virtualization (16 min)
- o Overview
- o Analysis
- o Design and Modeling
- o Development
- o Deployment and Operation
- o Review

Getting Started with Data Virtualization (28 min)
- o Overview
- o Skills, Competencies, and Human Factors
- o Goal and Expectations
- o Best Practices
- o Case Studies

OUR INSTRUCTORS

Andy Hayler

Andy Hayler is one of the world's foremost experts on master data management. Andy founded Kalido, which under his leadership was the fastest growing business intelligence vendor in the world in 2001. Andy was the only European named in Red Herring's "Top 10 Innovators of 2002". Kalido was a pioneer in modern data warehousing and master data management. He is now founder and CEO of The Information Difference, a boutique analyst and market research firm, advising corporations, venture capital firms and software companies. He is a regular keynote speaker at international conferences on master data management, data governance and data quality.

Kathy Hunter

Kathy always says she has data in her blood. Joining Harte-Hanks in 2002, she built an information management practice and, with her highly skilled team, was responsible for instituting their highly successful Global Data Management solution set. From information quality and data governance through to providing global data solutions and guidance she attained a reputation for expert knowledge and successful delivery in global information management to her clients. Kathy is known for her pragmatic approach to topics, providing helpful hints and practical examples in order to solve tough problems.

Theresa Kushner

Theresa Kushner is presently the Vice President of Enterprise Information Management for VMware, Palo Alto. She joined in October 2012 to help the fast growing software company develop a firm data foundation on which to build their future business. Before joining VMware she was the Director of Customer Intelligence within the Strategic Marketing organization of Cisco Systems.

William McKnight

William is president of McKnight Consulting Group, which includes service lines of Master Data Management, IT assessment, Big Data, Columnar Databases, Data Warehousing, and Business Intelligence. He functions as Strategist, Lead Enterprise Information Architect, and Program Manager for sites worldwide. Many of his clients have gone public with their success stories. William is a Southwest Entrepreneur of the Year Finalist, a frequent best practices judge, has authored hundreds of articles and white papers and given hundreds of international keynotes and public seminars. His team's implementations from both IT and consultant positions have won Best Practices awards. William is a former Information Technology Vice President of 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. As a practitioner, author and educator he has been involved in some of the most challenging projects industry has seen. These projects were often the result of major corporate mergers and the need to consolidate and integrate databases of enormous variety and complexity. Arkady's client list includes such household names as Dun & Bradstreet, Hewitt Associates, Kimberly Clark, Raytheon, Sprint, Verizon, and Xerox.

Olga Maydanchik

Olga Maydanchik is an experienced practitioner and educator in the field of Information Management. As a part of Chief Data Offices in Citi, AIG, Deutsche Bank, and Voya Financial Olga was focusing 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 workstreams.

Henrik Sørensen

Henrik Liliendahl Sørensen has over 30 years of experience in 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. Henrik is the founder of the Data Matching and the Multi-Domain MDM groups on LinkedIn.

Maria C. Villar

Maria C. Villar is a leader, consultant and writer in the field of enterprise information management, IT management and software development. She has held senior executive positions in both the technology and financial sector. Maria holds a bachelor in Computer Science and graduate degrees in Management Information Systems and Business Administration. Maria has guest lectured on the topic of IT and information management in leading universities, industry conferences and Fortune 500 companies across the country.

Dave 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 relationships of business and technology. This perspective —refined through a career of more than thirty-five years that encompassed both business and technical roles— helps to 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.

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

CONTACT US

Director, Enterprise Solutions
Arkady Maydanchik
Arkadym@elearningcurve.com

Director, Marketing
Michelle Johnson
Mjohnson@elearningcurve.com

Director, Education
Dave Wells
David.Wells@elearningcurve.com

Director, Technology
Varya Belyaevskaya
Admin@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

MDM COURSE PRICING

Packages

CIMP Master Data Management Package	$1,995.00
CIMP Ex Master Data Management Package	$2,995.00
CIMP Ex Data Quality & MDM Package	$3,595.00
CIMP Ex Data Integration & MDM Package	$3,495.00
CIMP All Course Access License	$4,995.00
CIMP All Course Access License w/ Exam Package	$5,990.00

Individual Courses

MDM Fundamentals: Architecture & Implementation	$460.00
Data Profiling	$520.00
Data Parsing, Matching & De-duplication	$435.00
Data Virtualization	$285.00
Data Governance Fundamentals	$500.00
Metadata Management Fundamentals	$385.00
Data Integration Fundamentals and Best Practice	$500.00
Data Quality Assessment	$510.00
Data Quality Scorecard	$515.00
Ensuring Data Quality in Data Integration	$525.00

Exams

CIMP Exam for each course….................................... $80.00

Enterprise Discounts

We offer discounts to Enterprise customers who purchase in bulk. Please contact us for more information.

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.