CDO-1 Certificate Program

Foundations for the Chief Data Officers

CDO-1 CERTIFICATE PROGRAM AGENDA

Day 1 (Virtual) June 19, 2023 (Monday)	Day 2 (On-site and Live Virtual) – **Attendees can pick either Track A, B, or C* July 17, 2023 (Monday)		
L1: 8:30 – 9:45 AM **Eastern Time** Introduction Richard Wang	**L7A: 8:30 – 10:15 AM** **Eastern Time** Data Governance, Part 1 Martha Dember	**L7B: 8:30 – 10:15 AM** **Eastern Time** Data Mesh - A Modern Decentralized Data Management Concept, Part 1 by Tamr	**L7C: 8:30 – 10:15 AM** **Eastern Time** Data Fabric, Part 1 by Modak
9:45 – 10:15 AM Break	**10:15 – 10:45 AM** Break		
L2: 10:15 – 11:30 AM What is CDO? Richard Wang	**L8A: 10:45 AM – 12:15 PM** Data Governance, Part 2 Martha Dember	**L8B: 10:45 AM – 12:15 PM** Data Mesh - A Modern Decentralized Data Management Concept, Part 2 by Tamr	**L8C: 10:45 AM – 12:15 PM** Data Fabric (Case Study), Part 2 by Modak
11:30 AM– 12:45 PM Lunch Break	**12:15 PM – 1:00 PM** Lunch Break		
L3: 12:45 – 2:00 PM A Cubic Framework for the CDO Yang Lee	**L9A: 1:00 – 2:45 PM** Data Quality Management, Part 1 Yang Lee	**L9B: 1:00 – 2:45 PM** Enabling Data Mesh for the Enterprise, Part 1 By PwC	**L9C: 1:00 – 2:45 PM** Experiences from implementing data mesh, Part 1 by Thoughtworks
2:00 – 2:30 PM Break	**2:45 – 3:15 PM** Break		
L4: 2:30 – 3:45 PM Data Policy & Strategy Richard Wang	**L10A: 3:15 – 4:45 PM** Data Quality Management, Part 2 Yang Lee	**L10B: 3:15 – 4:45 PM** Enabling Data Mesh for the Enterprise, Part 2 By PwC	**L10C: 3:15 – 4:45 PM** Experiences from implementing data mesh, Part 2 by Thoughtworks
3:45 – 4:15 PM Break	**4:45 – 5:15 PM** Break		
L5: 4:15 – 5:30 PM Panel Discussion: CDOs in Action Randy Bean, Carl Gerber, Mark Ramsey, Derek Strauss, and Maria Villar	**L11: 5:15 – 6:00 PM** **Selected Course Project Presentations** All Attendees & Richard Wang		
L6: 5:30 – 6:00 PM **Course Project Discussion** All Attendees & Richard Wang	**L12: 6:30 – 7:00 PM** **Conclusion, Certificate Ceremony & Welcome Party on 14th Floor** All Attendees & Richard Wang		

Co-Instructors

Martha Dember

Data Management Professional

Yang Lee

Associate Professor
Faculty Director,
Northeastern University

Richard Wang

Executive Director
@ CDOIQ Program

Randy Bean

Innovation Fellow @ Wavestone

Carl Gerber

CDO @ KPMG

Mark Ramsey

Managing Partner
@ Ramsey International

Derek Strauss

Co-founder @ Gavroshe

Maria Villar

Head of Enterprise Data Strategy & Transformation, SAP

LECTURE OVERVIEW

Day 1

Sessions	Description
L1: Introduction	The CDO-1 Certificate Program offered by Dr. Richard Wang will explain the definition and landscape of the Chief Data Officer, present data policy, data strategy, data governance, data analytics, data integration, and tools that are immediately deployable by CDOs. The certificate program will also include a required course project that attendees must complete during the one-month course period that can be applied to their enterprise upon completion of the course.
L2: What is CDO?	CDOs deliver business value and clean data to the enterprise. We broadly define CDO as the most senior executive who leads the data vision, mission, and culture across the organization with enterprise-wide governance to utilize information as an asset and deliver data and analytics capabilities in support of business strategy.
L3: A Cubic Framework for the Chief Data Officer	The rise of the chief data officer (CDO) is now well accepted by the government and industries. In this lecture, we present a three-dimensional cubic framework that describes the role of the CDO. The three dimensions are: (1) Collaboration Direction (inwards vs. outwards), (2) Data Space (traditional data vs. big data) and (3) Value Impact (service vs. strategy). We illustrate the framework with examples from early adopters of the CDO role and provide recommendations to help organizations assess and strategize the establishment of their own CDOs.
L4: Data Policy and Strategy	To make use of data effectively and efficiently, organizations must establish their own data policies and strategies. Data policy consists of the laws, rules, and norms that govern how data are collected, shared, stored, protected, accessed, analyzed, and disseminated. The changes in policies are typically driven by emerging data demands, evolving data sources, advances in technology, and organizational changes. The main purpose of having a data policy is to provide consistency and a set of governing rules for data that all members involved in the data enterprise can follow.
L5: CDOs in Action Panel Discussion	Dr. Wang will invite prominent data leaders who are current or former CDOs for this interactive session to discuss with attendees about their experiences including success stories and lessons learned.
L6: Course Project Discussion	The CDO-1 course project will benefit the students to develop a roadmap for a data practice during the project period. Students are encouraged to identify new problems/tasks/applications. A virtual class on June 19th will be held for the course project. During this meeting, the course instructor, Dr. Rich Wang will provide an overview of the course & interact with the students to discuss their topics.

Day 2

Sessions	Description
L7A & L8A: Data Governance	Data Governance: Past, Present and Future Data Governance sprang up in response to the needs of quality data highlighted by the implementation of data warehouses. As data lakes, snowflake models and the move to the cloud for example have changed the way data is stored and accessed so to has the way we manage data. This session will begin with understand how and why Data Governance came into existence with an understanding of what worked and what failed over the past 20 years. We will then take a deep dive into where things stand today and end with a glimpse into the future. Attendees will take away from this session: 1. An understanding of the core principles that have not changed, standing the test of time. 2. Knowledge of what it takes to sustain a data governance program for the long term. 3. Insight to the trends impacting where the future of data governance is heading.
L9A & L10A: Data Quality Management	We identify and analyze ten root conditions that can develop into data quality problems. We then propose intervention actions that organizations can take to prevent or redirect the negative progression of the root conditions and improve data quality.
L7B & L8B: Mastering Data Mesh: A Practical Guide to Decentralized Data Architecture by Tamr	Data mesh is a novel approach to designing and operating data systems in large and complex organizations. It advocates for treating data as a product and building a federated data architecture that enables decentralized ownership and governance of data domains. This tutorial provides an in-depth introduction to data mesh, its business applications, and its architecture principles. We will discuss the core concepts of data mesh, such as domain-driven design, bounded contexts, and data products, and explore how they can help organizations improve data quality, accelerate innovation, and foster data-driven decision-making. The tutorial will also cover practical aspects of implementing data mesh, such as designing data platforms and data pipelines, building data teams, and establishing data governance. We will delve into the role of data mastering in data mesh and how it can help organizations achieve a unified view of their data assets while maintaining domain autonomy. Finally, we will showcase a real-world case study of a company that is implementing a data mesh strategy and the benefits it gained from it. By the end of this tutorial, you will have a solid understanding of data mesh and be equipped with practical knowledge to start applying it to your own organization's data challenges.
L9B & L10B: Enabling Data Mesh for the Enterprise by PwC	Abstract: Data mesh is an architectural approach for managing and distributing data ownership across an organization. The goal is to improve the agility and responsiveness in delivering business outcomes for an organization by making data more accessible to domain-specific consumers. This approach involves a decentralized model for data ownership, distribution, and governance, with each domain managing its own data pipelines. Data mesh also encourages the use of a common data interoperability layer and standards to ensure data consistency, reliability and improves data literacy within the organization Learning Objectives:

	After completing this topic, learners will be able to understand the key principles of data mesh, including data ownership by domain, treating data as a product, self-serve data infrastructure as a platform, and federated data governance by domain. They will also learn about the benefits of using a data mesh approach, such as increased data accessibility and improved data consistency and reliability. Additionally, learners will understand the technical and organizational challenges involved in implementing data mesh, and they will be able to identify potential use cases for data mesh in their own organization. Overall, learners will be equipped with a comprehensive understanding of the data mesh approach and its potential to transform how organizations manage and distribute data, as well as how they can communicate internally to shape a business case for the same.
L7C & L8C: Data Fabric by Modak	Data fabric is a distributed architecture that connects various data sources and enables seamless data access across different systems, locations, and formats. It provides a unified view of data, regardless of where it resides, and enables organizations to harness the power of their data to make informed decisions. The Modak Data Fabric course is designed to provide individuals with a comprehensive understanding of data fabric technology and its applications. The purpose of this course is to enable data practitioners and business technologists to learn how to manage, integrate, and analyze data from multiple sources and locations. The course will provide a comprehensive understanding of the fundamentals of data fabric technology, its architecture, and its components. The audience will learn how to build data fabrics that can handle large volumes of data, integrate data from multiple sources, and perform real-time analytics. The course covers topics such as data integration, data governance, and self-service data management. The real-world examples of how Modak's implementation of enterprise data fabric is empowering leading organizations to achieve their data and business initiatives faster.
L9C & L10C: Experiences from implementing data mesh by Thoughtworks	This comprehensive tutorial designed and presented by Thoughtworks, is offered in three parts to give you Thoughtworks distilled experience of applying data mesh in different contexts and locations. What is data mesh: This section gives you a thorough understanding of the principles of Data Mesh and how they get applied. How these principles are defined and how they interact with each other. We will also discuss what are the motivations and objectives of Data Mesh and assess if they apply to your organization, compare data mesh with other data architectures. How to implement data mesh: This section gives you a deeper understanding of the architectural components of a Data Mesh implementation and their design. You will be presented with architectural concepts, design. Discuss Logical architecture of multi-plane data mesh platform, Logical architecture of data products including input ports and output ports. Discuss how to achieve federated computational governance and components such as discoverability and observability. We will discuss, organizational structures, transformation and walk you through models of implementation such as show->shift->scale What problems will you run into: This section we will discuss the lessons we have learned over the course of many years of implementing data mesh, the many obstacles the implement efforts encounter, either in the technology implementation phase, the people transformation, finding the people to fulfill the new job roles needed to adopt data mesh.

L11: Course Project Presentation	Each student is required to write up a 500-word overview of his/her course project. We encourage the students to work together if from the same company. Submit your course project overview to CDOIQ (cdoiq@mit.edu) by June 30th. Prepare a PPT before going to the onsite course and present your approach on July 17th.