

# DATA SCIENCE GOVERNANCE

## Training & Certification Guide



Arcitura<sup>®</sup>  
DIGITAL

# About the Next-Gen Data Science Academy

The Next-Gen Data Science Academy from Arcitura provides formal education and accreditation programs dedicated to the fields of Artificial Intelligence, Machine Learning, Big Data and general Data Science, including analytics and analysis, architecture, engineering and governance.

The Next-Gen Data Science Academy curriculum is comprised of 24 course modules and 9 certification tracks. Exams are available worldwide via Pearson VUE testing centers, as well as via Pearson VUE OnVue online proctoring and on-site delivery by Certified Trainers. Achieving a passing grade on the required exam(s) achieves a certification for which a digital accreditation certificate is automatically issued by Arcitura and a digital certification badge is issued by Acclaim/Credly.

## TABLE OF CONTENTS

03

How to Take This Course

04

How to Get Started

05

How to Get Certified

06

Course Module Outlines

10

Training and Exam Preparation Resources

11

Arcitura Certification Programs



## HOW TO TAKE THIS COURSE



### ARCITURA ELEARNING OPTIONS

To give you the most flexibility to achieve your learning goals and accommodate your preferences, this course is made available via two Arcitura eLearning solutions: An interactive environment with graded exercises and a graded self-test, as well as a study kit account that supports online/offline access and custom annotations.

To learn more, visit: [www.arcitura.com/elearning](http://www.arcitura.com/elearning)

To enroll, visit: [digital.arcitura.com/courses](http://digital.arcitura.com/courses)



### ONSITE WORKSHOPS

Private onsite workshops can be delivered by Certified Trainers from Arcitura and authorized training partners for both small and large groups. Each workshop participant can receive access to the full Arcitura Digital eLearning course materials. Each workshop agenda, format and schedule can be tailored to client requirements.

To learn more, visit: [www.arcitura.com/workshops](http://www.arcitura.com/workshops)



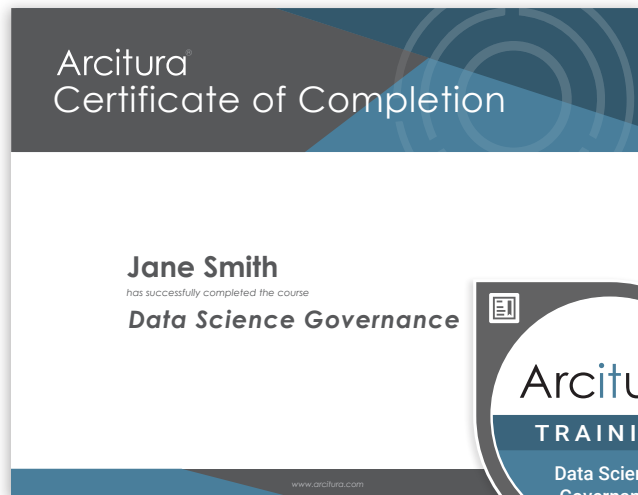
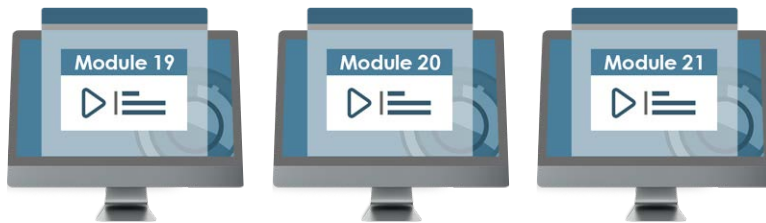
### VIRTUAL WORKSHOPS

Private virtual workshops can be delivered by Certified Trainers for small and large groups, as well as individual participants. Workshop participants receive access to the course materials via the Arcitura Digital eLearning platform. Virtual workshop agendas can be tailored with greater flexibility to accommodate more distributed and fragmented training schedules.

To learn more, visit: [www.arcitura.com/workshops](http://www.arcitura.com/workshops)



Several additional learning and exam preparation products and services are available, including coaching, exam prep kits and digital downloads. See the Training and Exam Preparation Resources page for details.



## HOW TO GET STARTED

Welcome to the Data Science Governance course! This course is comprised of a set of modules. Each module has a set of lessons and is further supplemented with exercises to help reinforce your understanding of key topics. Upon completing the course, you can optionally proceed to prepare yourself for the certification exam (as explained on the *How to get Certified* page).

Additional resources are available to assist you with completing this course, including downloadable digital course files, printed course materials, coaching hours and instructor-led training services (as explained on the *Training and Exam Preparation Resources* page.)

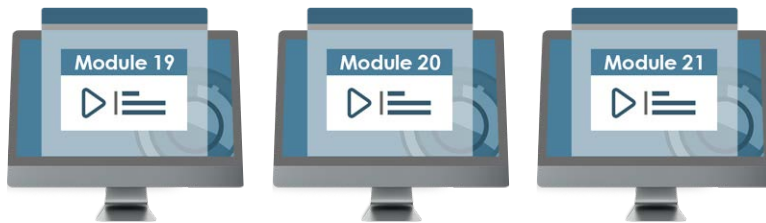
## PREREQUISITE

The prerequisite for this course is the completion of the Big Data Analytics & Fundamental Data Science course. It is recommended that you complete the Big Data Analytics & Fundamental Data Science course prior to starting the Data Science Governance course.

### MORE INFO

To enroll in this course, visit:  
[digital.arcitura.com/courses](https://digital.arcitura.com/courses)





## HOW TO GET CERTIFIED

You can become a Certified Data Science Governance Specialist! This course can prepare you for the official Data Science Governance Specialist Certification exam, which can be taken worldwide at Pearson VUE testing centers, via Pearson VUE online proctoring and/or Arcitura direct proctoring.

Upon attaining a passing grade on the certification exam (and fulfilling any other prerequisite exam requirements) you will achieve the Data Science Governance Specialist Certification, after which you will automatically receive an official digital Accreditation Certificate and a digital Certification Badge from Acclaim/Credly with an account that supports the online verification of your certification status. Digital accreditation certificates and badges are free of charge.

Additional resources are available to assist you with preparing for the certification exam, including practice exam questions, downloadable digital course files, printed course materials, coaching hours and instructor-led training services (as explained on the *Training and Exam Preparation Resources* page.)

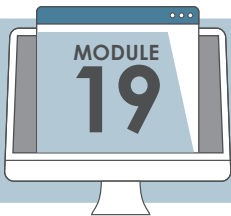
### MORE INFO

To learn more about this certification and exam, visit:  
[www.arcitura.com/certifications](http://www.arcitura.com/certifications)

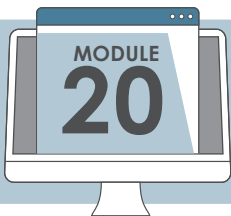
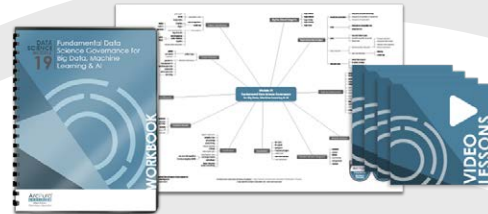


## COURSE MODULE OUTLINES

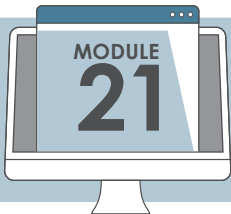
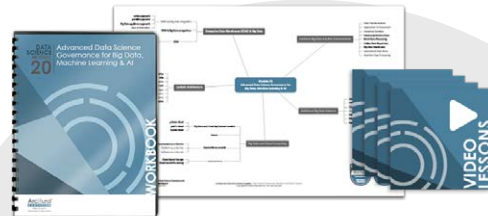
The Big Data Governance course is comprised of the following course modules. Outlines for these course modules are provided on the subsequent pages.



Fundamental Data Science Governance for Big Data, Machine Learning & AI



Advanced Data Science Governance for Big Data, Machine Learning & AI

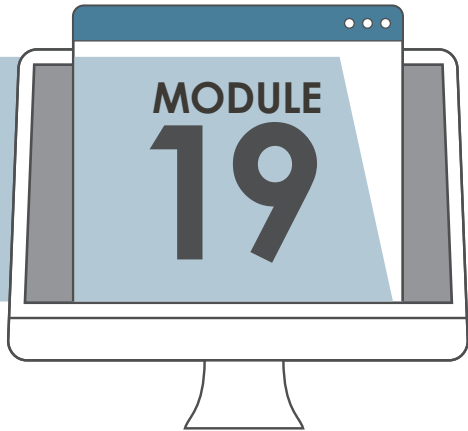


Data Science Governance Lab for Big Data, Machine Learning & AI



# Fundamental Data Science

## Governance for Big Data, Machine Learning & AI

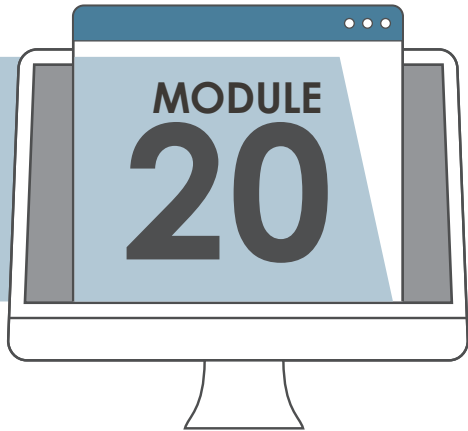


This course module explores introductory topics pertaining to the field of developing data processing solutions—data engineering—in the context of Big Data environments. Specifically it covers concepts, techniques and technologies related to the processing and storage of Big Data datasets including MapReduce and NoSQL. It highlights the unique challenges faced when processing and storing Big Data datasets. The MapReduce data processing engine, which is the de facto framework for batch processing of large amounts of data, is also explained in detail.

The following primary topics are covered:

- Big Data Engineering – Big Data Engineering Challenges
- Big Data Storage Terminologies (including sharding, replication, CAP theorem, ACID, BASE)
- Big Data Storage Requirements
- On-Disk Storage (including distributed file system – databases)
- Introduction to NoSQL – NewSQL
- NoSQL Rationale – Characteristics
- NoSQL Database Types (including key-value, document, column-family and graph databases)
- Big Data Processing Requirements
- Big Data Processing (including batch mode and realtime mode)
- Introduction to MapReduce for Big Data Processing (batch mode)
- MapReduce Explained (including map, combine, partition, shuffle and sort, and reduce)

# Advanced Data Science Governance for Big Data, Machine Learning & AI



//////

This course module builds upon Module 7 by exploring advanced topics pertaining to the storage and processing of Big Data datasets. Specifically it covers advanced Big Data engineering mechanisms, in-memory data storage and realtime data processing. It presents further considerations for developing MapReduce algorithms and also introduces the Bulk Synchronous Parallel (BSP) processing engine, along with a discussion of graph data processing. The Big Data mechanisms required for developing Big Data pipelines, its stages and the design process involved in developing Big Data processing solutions are also explored.

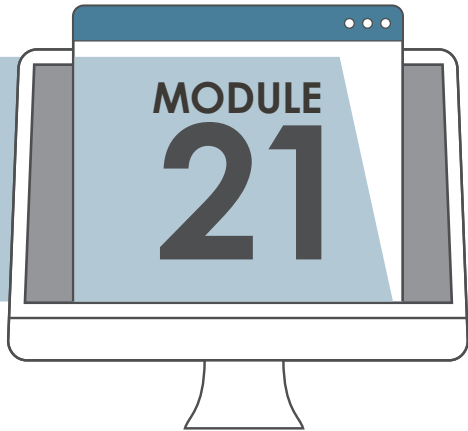
The following primary topics are covered:

- Advanced Big Data Engineering Mechanisms (including serialization & compression engines)
- In-Memory Storage Devices, In-Memory Data Grids & In-Memory Databases
- Read-Through, Read-Ahead, Write-Through & Write-Behind Integration Approaches
- Polyglot Persistence (including Explanation, Issues & Recommendations)
- Realtime Big Data Processing Concepts (including Speed Consistency Volume (SCV), Event Stream Processing (ESP) & Complex Event Processing (CEP))
- General Realtime Big Data Processing & Realtime Big Data Processing & MapReduce
- Advanced MapReduce Algorithm Design
- Bulk Synchronous Parallel (BSP) Processing Engine & BSP vs. MapReduce
- Graph Data & Graph Data Processing using BSP
- Big Data Pipelines (including Definition and Stages)
- Big Data with Extract-Load-Transform (ELT)
- Big Data Solutions (including Characteristics, Design Considerations & Design Process)



# Data Science Governance

## Lab for Big Data, Machine Learning & AI



This course module covers a series of exercises and problems designed to test the participant's ability to apply knowledge of topics covered previously in course modules 19 and 20. Completing this lab will help highlight areas that require further attention, and will further prove hands-on proficiency in Big Data engineering practices as they are applied and combined to solve real-world problems.

As a hands-on lab, this module incorporates a set of detailed exercises that require participants to solve various inter-related problems, with the goal of fostering a comprehensive understanding of how different data engineering technologies, mechanisms and techniques can be applied to solve problems in Big Data environments.

For instructor-led delivery of this lab module, the Certified Trainer works closely with participants to ensure that all exercises are carried out completely and accurately. Attendees can voluntarily have exercises reviewed and graded as part of the class completion. For individual completion of this course as part of the Module 9 Study Kit, a number of supplements are provided to help participants carry out exercises with guidance and numerous resource references.



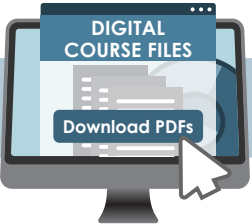
# TRAINING AND EXAM PREPARATION RESOURCES

You can supplement this course with a number of available resources to assist with both learning and exam preparation. Contact [info@arcitura.com](mailto:info@arcitura.com) with any questions.



## Certification Exam Prep Kit

A set of additional practice questions is available to support exam preparation.



## Digital Course Files

For each course you can order a set of downloadable digital course materials comprised of printable, watermarked workbook and poster PDF files.



## Printed Course Materials

The printed workbooks and posters for each course can be ordered in B&W and full-color, and can be shipped worldwide.



## One-on-One Coaching

Certified Trainers are available to provide online coaching on an hourly basis and in all time zones.



## Instructor-Led Training

Certified Trainers are available to provide virtual and onsite training workshops for this and other Arcitura courses.

COURSES		Digital Transformation	Digital Transformation: Fundamental Technology	Digital Transformation: Advanced Technology & Architecture	Digital Transformation: Fundamental Data Science	Digital Transformation: Advanced Data Science	Digital Transformation: Fundamental Security	Digital Transformation: Advanced Security	Digital Transformation: Fundamental Intelligent Automation	Digital Transformation: Advanced Intelligent Automation
CERTIFICATIONS		Digital Transformation Specialist	Digital Transformation Technology Professional	Digital Transformation Technology Architect	Digital Transformation Data Science Professional	Digital Transformation Data Scientist	Digital Transformation Security Professional	Digital Transformation Security Specialist	Digital Transformation Intelligent Automation Professional	Digital Transformation Intelligent Automation Specialist
MODULE 01	Fundamental Digital Transformation	●	●	●	●	●	●	●	●	●
MODULE 02	Digital Transformation in Practice	●	●	●	●	●	●	●	●	●
MODULE 03	Fundamental Cloud Computing		●	●						
MODULE 04	Fundamental Blockchain		●	●			●	●		
MODULE 05	Fundamental IoT		●	●						
MODULE 06	Cloud Architecture			●						
MODULE 07	Blockchain Architecture			●				●		
MODULE 08	IoT Architecture			●						
MODULE 09	Fundamental Big Data Analysis & Analytics				●	●				
MODULE 10	Fundamental Machine Learning				●	●				
MODULE 11	Fundamental AI				●	●				
MODULE 12	Advanced Big Data Analysis & Analytics					●				
MODULE 13	Advanced Machine Learning					●				
MODULE 14	Advanced AI					●				
MODULE 15	Fundamental Cybersecurity						●	●		
MODULE 16	Advanced Cybersecurity							●		
MODULE 17	Fundamental RPA								●	●
MODULE 18	Advanced RPA & Intelligent Automation									●
MODULE 19	Fundamental AI Decisioning								●	●
MODULE 20	Advanced AI Decisioning									●

Gray circles indicate prerequisite modules.

COURSES		DevOps	Blockchain Architecture	IoT Architecture	Cybersecurity	Robotic Process Automation	Digital Business Technology	Containerization Architecture	Quantum Computing
CERTIFICATIONS		DevOps Specialist	Blockchain Architect	IoT Architect	Cybersecurity Specialist	RPA Specialist	Digital Business Technology Professional	Containerization Architect	Quantum Computing Specialist
DevOps	MODULE 01 Fundamental DevOps	●							
	MODULE 02 DevOps in Practice	●							
	MODULE 03 DevOps Lab	●							
Blockchain	MODULE 01 Fundamental Blockchain		●						
	MODULE 02 Blockchain Technology & Architecture		●						
	MODULE 03 Blockchain Technology & Architecture Lab		●						
Internet of Things	MODULE 01 Fundamental IoT			●					
	MODULE 02 IoT Technology & Architecture			●					
	MODULE 03 IoT Technology & Architecture Lab			●					
Cybersecurity	MODULE 01 Fundamental Cybersecurity				●				
	MODULE 02 Advanced Cybersecurity				●				
	MODULE 03 Cybersecurity Lab				●				
Robotic Process Automation	MODULE 01 Fundamental RPA					●			
	MODULE 02 Advanced RPA & Intelligent Automation					●			
	MODULE 03 RPA Lab					●			
Digital Business Technology	MODULE 01 Business Automation Technology Overview						●		
	MODULE 02 Data Science Technology Overview						●		
	MODULE 03 Digital & Security Technology Overview						●		
Containerization	MODULE 01 Fundamental Containerization							●	
	MODULE 02 Containerization Technology & Architecture							●	
	MODULE 03 Containerization Technology & Architecture Lab							●	
Quantum Computing	MODULE 01 Fundamental Quantum Computing								●
	MODULE 02 Advanced Quantum Computing								●
	MODULE 03 Quantum Computing Lab								●

COURSES		Big Data Analytics & Fundamental Data Science	Big Data Analysis & Advanced Data Science	Data Science Professional Consulting	Machine Learning	Artificial Intelligence	Big Data Engineering	Big Data Architecture	Data Science Governance	AI Decisioning
CERTIFICATIONS		Big Data Science Professional	Big Data Scientist	Data Science Consultant	Machine Learning Specialist	Artificial Intelligence Specialist	Big Data Engineer	Big Data Architect	Data Science Governance Specialist	AI Decisioning Specialist
MODULE 01	Fundamental Big Data Science & Analytics	●	●	●			●	●	●	●
MODULE 02	Big Data Analysis & Technology Concepts	●	●	●			●	●	●	●
MODULE 03	Big Data Analysis & Technology Lab	●		●						
MODULE 04	Big Data Analysis & Science		●							
MODULE 05	Advanced Big Data Analysis & Science		●							
MODULE 06	Big Data Analysis & Science Lab		●							
MODULE 07	Fundamental Machine Learning			●	●					
MODULE 08	Advanced Machine Learning				●					
MODULE 09	Machine Learning Lab				●					
MODULE 10	Fundamental Artificial Intelligence			●		●				
MODULE 11	Advanced Artificial Intelligence					●				
MODULE 12	Artificial Intelligence Lab					●				
MODULE 13	Fundamental Big Data Engineering						●			
MODULE 14	Advanced Big Data Engineering						●			
MODULE 15	Big Data Engineering Lab						●			
MODULE 16	Fundamental Big Data Architecture							●		
MODULE 17	Advanced Big Data Architecture							●		
MODULE 18	Big Data Architecture Lab							●		
MODULE 19	Fundamental Data Science Governance for Big Data, Machine Learning & AI								●	
MODULE 20	Advanced Data Science Governance for Big Data, Machine Learning & AI								●	
MODULE 21	Data Science Governance Lab for Big Data, Machine Learning & AI								●	
MODULE 22	Fundamental AI Decisioning									●
MODULE 23	Advanced AI Decisioning									●
MODULE 24	AI Decisioning Lab									●

Gray circles indicate prerequisite modules.



COURSES		Cloud Computing	Cloud Computing Professional Consulting	Cloud Architecture	Cloud Security	Cloud Governance	Cloud Storage	Cloud Virtualization
CERTIFICATIONS		Cloud Technology Professional	Cloud Computing Consultant	Cloud Architect	Cloud Security Specialist	Cloud Governance Specialist	Cloud Storage Specialist	Cloud Virtualization Specialist
MODULE 01	Fundamental Cloud Computing	●	●	●	●	●	●	●
MODULE 02	Cloud Technology Concepts	●	●	●	●	●	●	●
MODULE 03	Cloud Technology Lab	●	●					
MODULE 04	Fundamental Cloud Architecture		●	●				
MODULE 05	Advanced Cloud Architecture			●				
MODULE 06	Cloud Architecture Lab			●				
MODULE 07	Fundamental Cloud Security		●		●			
MODULE 08	Advanced Cloud Security				●			
MODULE 09	Cloud Security Lab				●			
MODULE 10	Fundamental Cloud Governance					●		
MODULE 11	Advanced Cloud Governance					●		
MODULE 12	Cloud Governance Lab					●		
MODULE 13	Fundamental Cloud Storage						●	
MODULE 14	Advanced Cloud Storage						●	
MODULE 15	Cloud Storage Lab						●	
MODULE 16	Fundamental Cloud Virtualization							●
MODULE 17	Advanced Cloud Virtualization							●
MODULE 18	Cloud Virtualization Lab							●

Gray circles indicate prerequisite modules.

COURSES		Fundamental Microservices & Service Technology	Fundamental SOA Design with Services & Microservices	SOA Analysis & Modeling with Services & Microservices	SOA Design & Architecture with Services & Microservices	Microservice Design & Architecture	Microservice Professional Consulting	Service API Design & Management	Service Governance & Project Delivery	Security for Microservices & SOA
CERTIFICATIONS		Microservice Professional	SOA Professional	SOA Analyst	SOA Architect	Microservice Architect	Microservice Consultant	Service API Specialist	Service Governance Specialist	Service Security Specialist
MODULE 01	Fundamental SOA, Services & Microservices	●	●	●	●	●	●	●	●	●
MODULE 02	Microservice Technology Concepts	●			●	●	●	●		●
MODULE 03	Design & Architecture with SOA, Services & Microservices		●	●	●				●	
MODULE 04	Fundamental SOA Analysis & Modeling with Services & Microservices			●						
MODULE 05	Advanced SOA Analysis & Modeling with Services & Microservices			●						
MODULE 06	SOA Analysis & Modeling Lab with Services & Microservices			●						
MODULE 07	Advanced SOA Design & Architecture with Services & Microservices				●					
MODULE 08	SOA Design & Architecture Lab with Services & Microservices				●					
MODULE 09	Fundamental Microservice Architecture & Containerization					●	●			
MODULE 10	Advanced Microservice Architecture & Containerization					●				
MODULE 11	Microservice Architecture & Containerization Lab					●				
MODULE 12	Fundamental Service API Design & Management						●	●		
MODULE 13	Advanced Service API Design & Management							●		
MODULE 14	Service API Design & Management Lab							●		
MODULE 15	Fundamental Service Governance & Project Delivery								●	
MODULE 16	Advanced Service Governance & Project Delivery								●	
MODULE 17	Service Governance & Project Delivery Lab								●	
MODULE 18	Fundamental Security for Services, Microservices & SOA						●			●
MODULE 19	Advanced Security for Services, Microservices & SOA									●
MODULE 20	Security Lab for Services, Microservices & SOA									●

Gray circles indicate prerequisite modules.



Arcitura®

Copyright © Arcitura Education Inc.  
[www.arcitura.com](http://www.arcitura.com) • [info@arcitura.com](mailto:info@arcitura.com)