

# Big Data Analytics

## Program Overview

The Big Data Analytics Certification is essential for everyone working with Big Data. This certification provides an understanding of Big Data, potential data sources that can be used to solve real world problems, and an overview of data mining and its capabilities. Certification will help you become a Big Data Ready Professional.

## Objective

Candidates can expect to gain knowledge and understanding in the following upon successful completion

of the education and examination components related to this certification.

- Explain Big Data, its origin, and its characteristics.
- Discuss about the tools applicable to Big Data processing.
- Explain data mining.
- Discuss the popular Big Data technologies – Hadoop and MongoDB.
- Discuss the Big Data projects and the main players involved.
- Identify and obtain relevant datasets when looking at a business problem.
- Install and manage Big Data processing environments based on Hadoop or MongoDB at a departmental level.

## Target

### Audience:

- Software Engineers
- Application Developers
- IT Architects
- System administrators

## Workshop Chronology

Day 1: 9am – 5pm

Day 2: 9am – 5pm

Day 3: 9am – 5pm

Day 4: 9am – 5pm

Day 5: 9am – 5pm

• Exam will be on Day 3 for 1 hour  
(Closed Book)

• There is 40 Multiple Choice Questions

• 65% is the Passing rate

**BECOME A BIG DATA PROFESSIONAL**

# Big Data Analytics

## Workshop Outline

### DAY 1

#### Module 1. Big Data Fundamentals

##### 1.1 Big Data – History, Overview, and Characteristics

- History
- Big Data Definition
- Big Data Benefits
- Big Data – Volume, Velocity & Variety Characteristics

##### 2. Big Data Technologies – Overview

##### 3. Big Data Success Stories

##### 4. Big Data – Privacy and Ethics

- Privacy – Compliance
- Privacy – Challenges
- Privacy – Approach
- Ethics

##### 1.5 Big Data Projects

- Who Should Be Involved?
- What Is Involved?

### DAY 2

#### Module 2. Big Data Sources

##### 2.1 Enterprise Data Sources

- Enterprise Systems
- Oracle
- SAP
- Microsoft
- Data Warehouses
- Unstructured Data – Introduction
- Unstructured Data – Metadata

##### 2.2 Social Media Data Sources

- Introduction
- Facebook – Introduction
- Facebook – Public Feed API

- Twitter – Streaming APIs
- Twitter – REST APIs
- Twitter – REST APIs
- Other Social Media
- 2.3 Public Data Sources
- Introduction
- Weather
- Economics
- Finance
- Regulatory Bodies

### DAY 3

#### Module 3. Data Mining – Concepts and Tools

##### 3.1 Data Mining – Introduction

- Introduction
- Types of Data Mining – Overview
- Types of Data Mining – Classification
- Types of Data Mining – Association
- Types of Data Mining – Clustering

##### 3.2 Data Mining – Tools

- Introduction
- Weka
- Modules of Weka Applications
- KNIME
- KNIME – Example
- R Language

# Big Data Analytics

## Workshop Outline

### **DAY 3 (Continue)**

#### **Module 4. Big Data Technologies – Hadoop**

##### **4.1 Hadoop Fundamentals**

- Introduction
- Main Components of Hadoop
- Additional Components of Hadoop

##### **4.2 Install and Configure**

- Download
- How to Install and Configure

##### **4.3 MapReduce**

- Introduction
- How Does It Work?

##### **4.4 Data Processing with Hadoop**

- Introduction
- Twitter Sentiment Analysis – Overview
- Twitter Sentiment Analysis – Algorithm
- Network Log Analysis – Overview
- Network Log Analysis – Algorithm

## **Program Certified By:**



**CCSD COUNCIL**  
ISO 9001:2015 certified company

### **DAY 4**

#### **Module 5. Big Data Technologies – MongoDB**

##### **5.1 MongoDB Fundamentals**

- Introduction
- Replication
- Sharding
- Sharding and Replication
- MongoDB Ecosystem – Languages and Drivers
- MongoDB Ecosystem – Hadoop Integration
- Network Log Analysis – Algorithm
- MongoDB Ecosystem – Tools

##### **5.2 Install and Configure**

- Download
- How to Install and Configure

##### **5.3 Document Databases**

- Introduction
- Twitter Sentiment Analysis
- Twitter Sentiment Analysis – Algorithm
- Network Log Analysis

##### **5.4 Data Modelling with Document Databases**

- Introduction
- Documents
- Document Design Considerations
- Fields

### **Examination**

**Shieh Management Sdn Bhd (830400-M)**

8-12-4, Menara Mutiara Bangsar, Jalan Liku Off Jalan Riong, 59100 Bangsar Kuala Lumpur

Mobile: +6019 601 1148

Email: [info@shiehmgmt.com](mailto:info@shiehmgmt.com)



# Big Data Analytics

## Trainer Profile

### **Vigneswaren Krishnamoorthy**

•Big Data Solutions • Analytics and Visualization • Machine Learning and Artificial Intelligence • Graph Database Solutions • Be-Spoke Software Solutions • Training and Consulting

Vigneswaren has been involved in software development for the past 24 years, from a junior developer, researcher till a Senior member of management in a multinational. He still maintains an active continuous learning practice picking up the latest technologies and applying them to solve problems in the industry. As a continuous learning advocate he encourages the staff and people around him to pick up new skills to be competitive in the industry.

He has successfully delivered large scale project implemented globally and locally including HRDF eTRIS as a key member to formulate transformation of their online services. His other projects have been in various other spaces including government, medical, telecommunications and finance. As part of the projects in recent times he has implemented cutting edge technologies that cover big data mining, analytics, machine learning and AI in the solutions.

He has provided trainings across the range from user software training to technical trainings including Big Data, Machine Learning and AI, Graph Database and Mind Mapping essentials. As a Solution Architect he works closely with clients to define scope and develops strategy to deliver on the requirements of the projects with development and deliver teams of the organization.

He currently provides consulting and training for companies who are looking to improve their software development practices and technology stack implementation and is actively involved in forums for technology and workforce improvement.



**Shieh Management Sdn Bhd (830400-M)**

8-12-4, Menara Mutiara Bangsar, Jalan Liku Off Jalan Riong, 59100 Bangsar  
Kuala Lumpur

Mobile: +6019 601 1148

Email: [info@shiehmgmt.com](mailto:info@shiehmgmt.com)