



**big
data**

Big Data & Analytics From Strategy to Implementation

OVERVIEW

This two day workshop is aimed at getting Data Scientists, Data Warehousing and BI professionals up to scratch on Big Data, Hadoop, other NoSQL DBMSs and Multi-Platform Analytics. What is Big Data? How can you make use of it? How does it fit within a traditional analytical environment? What skills do you need to develop for Big Data Analytics? All of these questions are addressed in this newknowledge packed workshop.

AUDIENCE

IT directors, CIO's, IT Managers, BI Managers, data warehousing professionals, data scientists, enterprise architects, data architects

LEARNING OBJECTIVES

Attendees to this seminar will learn:

- ✓ What Big Data is
- ✓ How Big Data creates several new types of analytical workload
- ✓ Big Data technology platforms beyond the data warehouse
- ✓ Big Data analytical techniques and front-end tools
- ✓ How to analyse un-modelled, multi-structured data using Hadoop, MapReduce & Spark
- ✓ How to integrate Big Data with traditional data warehouses and BI systems
- ✓ How to clearly understand business use cases for different Big Data technologies
- ✓ How to set up and organise Big Data projects including skills
- ✓ How to make use of Big Data to deliver business value

MODULE 1:

AN INTRODUCTION TO BIG DATA

This session defines big data and looks at business reasons for wanting to make use of this new area of technology. It looks at Big Data use cases and what the difference is between traditional BI and Data Warehousing versus Big Data

- ✓ What is Big Data?
- ✓ Types of Big Data
- ✓ Why analyse Big Data?
- ✓ The need to analyse new more complex data sources
- ✓ Industry use cases - Popular big data analytic applications
- ✓ What is Data Science?
- ✓ Data Warehousing and BI Versus Big Data
- ✓ Popular patterns for Big Data technologies



Big Data & Analytics - From Strategy to Implementation

MODULE 2:



AN INTRODUCTION TO BIG DATA ANALYTICS

This session looks at Big Data Analytical workloads, the technology components involved and how you can integrate these with existing DW/BI systems in a new architecture for end-to-end analytics and to enrich business insight. It also looks at how to preserve existing investment in data management and BI tools across DW and Big Data platforms

- ✓ Types of Big Data analytical workloads
- ✓ Streaming data analytics at high velocity
- ✓ Exploratory analysis of multi-structured data
- ✓ Complex analysis of structured data
- ✓ Graph analytics
- ✓ Challenges when managing and analysing big data
- ✓ Key components in a Big Data Analytics environment
- ✓ Preserving existing BI/DW investments
- ✓ The Big Data Extended Analytical Ecosystem

MODULE 3:



BIG DATA PLATFORMS AND STORAGE OPTIONS

This session looks at platforms and data storage options for big data analytics

- ✓ The new multi-platform analytical ecosystem
- ✓ Beyond the data warehouse - Hadoop NoSQL and analytical RDBMSs, NewSQL DBMSs
- ✓ NoSQL DBMSs
 - Key Value stores, Document DBMSs, Column Family DBMSs and Graph databases
- ✓ An introduction to Hadoop and the Hadoop Stack
- ✓ HDFS, MapReduce, Pig & Hive
- ✓ Apache Spark Framework
- ✓ SQL on Hadoop options
 - Impala, Hive, SparkSQL, HawQ, HP Vertica SQL on Hadoop, IBM BigSQL, CitusData, Jethro, Splice Machine, ActianAnalytics Platform, Oracle Big Data SQL, Teradata QueryGrid
- ✓ The Big Data Marketplace
 - Hadoop distributions - Cloudera, Hortonworks, MapR, BM Big Insights, IBM Open Platform with Apache Hadoop, Microsoft HD Insight
 - Big Data Appliances - Oracle Big Data Appliance, IBM PureData System for Hadoop, HPE Big Data Platform, Teradata Aster Discovery Server
 - NoSQL databases, e.g. DataStax, Neo4j, Cray, MongoDB, Basho Riak
- ✓ The Cloud deployment option - Microsoft Azure (HDInsight, Data Lake + Data Factory) IBM Bluemix, Amazon Elastic MapReduce, Altiscale (a SAP Company) Data Cloud, Qubole, Oracle Analytics Cloud
- ✓ Creating a multi-platform analytical ecosystem



Big Data & Analytics - From Strategy to Implementation

MODULE 4:



BIG DATA INTEGRATION AND GOVERNANCE IN A MULTI-PLATFORM ANALYTICAL ENVIRONMENT

This session will look at the challenge of integrating and governing Big Data and the unique issues it raises. How do you deal with very large data volumes and different varieties of data? How does loading data into Hadoop differ from loading data into analytical relational databases? What about NoSQL databases? How should low-latency data be handled? Topics that will be covered include:

- ✓ Types of Big Data
- ✓ Connecting to Big Data sources, e.g. web logs, clickstream, sensor data, and multi-structured content
- ✓ Supplying consistent data to multiple analytical platforms
- ✓ Loading Big Data – what's different about loading HDFS, Hive & NoSQL Vs analytical relational databases
- ✓ Change data capture – what's possible
- ✓ Data warehouse offload
- ✓ Tools for ELT processing on Hadoop – The Enterprise Data Refinery
- ✓ ETL tools Vs Pig Vs self-service DI/DQ
- ✓ Dealing with data quality in a Big Data environment
- ✓ Parsing unstructured data
- ✓ Governing data in a Data Science environment
- ✓ Joined up analytical processing from ETL to analytical workflows
- ✓ The impact of data scientist and end user self-service DQ/DI – Paxata, Trifacta, MS Excel Power Query, MicroStrategy, Tableau
- ✓ Mapping discovered data of value into your DW and business vocabulary
- ✓ Big data audit, protection and security -Cloudera Sentry, Dataguise, Hortonworks Ranger, IBM Security Guardium, Protegrity

MODULE 5:



TOOLS AND TECHNIQUES FOR ANALYSING BIG DATA

This session looks at platforms and data storage options for big data analytics

- ✓ Data Science projects
- ✓ Creating Sandboxes for Data Science projects
- ✓ Options for analysing unstructured content – Text analytics, custom MapReduce code and MapReduce developer tools
- ✓ Using R as an analytical language for Big Data
- ✓ Text analysis and visualisation, Sentiment analysis and visualisation
- ✓ Clickstream analysis and visualisation
- ✓ Analysing big data using MapReduce BI Tools and applications for Hadoop, e.g. ClearStory Data, Datameer, FICO Big Data Analyser
- ✓ Exploratory graph analysis and visualisations
- ✓ Using search to analyse multi-structured data
 - Creating search indexes on multi-structured data
 - Building dashboards and reports on top of search engine indexed content
 - The integration of search with traditional BI platforms
 - Guided analysis using multi-faceted search
 - The marketplace: Apache Solr, Attivio, Cloudera Search, Connexica, DataRPM, HPE IDOL, IBI WebFocus Magnify, IBM Watson Explorer, LucidWorks, Microsoft, Oracle Endeca, Oracle Big Data Discovery, Quid, Splunk
- ✓ Analysing Big Data using Self-Service BI/Analytics Tools, e.g. Dell Statistica, Excel, IBM Watson Analytics, MicroStrategy, Qlik, RapidMiner, SAP BusinessObjectsLumira, SAS Visual Analytics, Tableau, TIBCO Spotfire, Zoomdata
- ✓ Big data analytics – query performance enablers
- ✓ Managing stream computing in a Big Data environment
- ✓ Tools and techniques for streaming analytics



Big Data & Analytics - From Strategy to Implementation

MODULE 6:



INTEGRATING BIG DATA ANALYTICS INTO THE ENTERPRISE

This session looks at how new Big Data platforms can be integrated with traditional Data Warehouses and Data Marts. It looks at stream processing, Hadoop, NoSQL databases, Data Warehouse appliances and shows how to put them together in an end-to-end architecture to maximise business value from Big Data

- ✓ Integrating Big Data platforms with traditional DW/BI environments – what's involved
- ✓ Integrating stream processing with Hadoop and Analytical DW Appliances
- ✓ Integrating Hadoop with DW Appliances and Enterprise Data Warehouses
- ✓ Tying together front end tools
- ✓ Options for implementing multi-platform analytics
- ✓ Cross-platform analytical workflows
- ✓ The role of Data Virtualisation in a Big Data environment
- ✓ Multi-platform optimisation

Who Should attend?



IT directors, CIO's, IT Managers, BI Managers, data warehousing professionals, data scientists, enterprise architects, data architects and others wanting to up to scratch on Big Data, Hadoop, other NoSQL DBMSs and Multi-Platform Analytics.

What are the learning objectives?



Attendees on this seminar will learn:

- ✓ What Big Data is
- ✓ How Big Data creates several new types of analytical workload
- ✓ Big Data technology platforms beyond the data warehouse
- ✓ Big Data analytical techniques and front-end tools
- ✓ How to analyse un-modelled, multi-structured data using Hadoop, MapReduce & Spark
- ✓ How to integrate Big Data with traditional data warehouses and BI systems
- ✓ How to clearly understand business use cases for different Big Data technologies
- ✓ How to set up and organise Big Data projects including skills
- ✓ How to make use of Big Data to deliver business value

PRESENTER



Mike Ferguson is Managing Director of Intelligent Business Strategies Limited. As an analyst and consultant he specialises in BI/Analytics, Big Data and Data Management. With over 34 years of IT experience, Mike has consulted for dozens of companies on BI, technology selection, Big Data, enterprise architecture, and data management. He has spoken at events all over the world and written numerous articles. Mike provides articles, blogs and his insights on the industry. Formerly he was a principal and co-founder of Codd and Date Europe Limited – the inventors of the Relational Model, a Chief Architect at Teradata on the Teradata DBMS and European Managing Director of Database Associates. He teaches popular master classes in BI, Big Data Analytics, Data Governance & Master Data Management.



Big Data & Analytics - From Strategy to Implementation

Ticket Prices

07 Dec 2016, London

2 Days Course

Price:

£1095.00 +VAT

Contact



UNICOM Seminars Ltd
OptiRisk R&D House
One Oxford Road
Uxbridge UB9 4DA, UNITED KINGDOM



+44 (0) 1895 256484



+44 (0) 1895 813095



www.unicom.co.uk



Info@unicom.co.uk