

Updated 05/28/2024

Sign up

Big Data training: State of the art

2 days (14 hours)

Presentation

Big Data is emerging as one of the major trends in information systems, thanks to the significant increase in the amount of data available to companies and public bodies, in terms of business, functions and technology. With so much potential already present within structures, it's no coincidence that it's one of the hottest topics for IT and business departments: but you still need to find your value in it! That's what our experts are here to help you discover.

From the Data Lake to the Data Lab, the world of Big Data is a rich and complex ecosystem that IT and Business Departments need to master if they are to support the data transformation of companies.

In this training course, you'll discover the state of the art in Big Data, so that you can set out a strategic vision for your company, thanks to concrete examples put into practice by our consultants from the field. In just two days, you'll know which directions to take and establish a roadmap for your teams over the next few years.

Objectives

- Understanding the main concepts of Big Data
- Identifying the benefits of Big Data
- Assess the impact on the company's organization and anticipate its integration
- Discover the advantages and limits of Big Data
- Understanding the economic challenges of Big Data
- Understand the Big Data ecosystem and associated technologies
- Consider data security and confidentiality

Target audience

Managers, Project Managers, Technical Directors, Project Leaders, Developers, Architects

Prerequisites

In-depth general IT culture.

BigData training program

Discovering where data can help

- The Data Strategy Document
- The Expected Value Analysis Framework
- The Confusion Matrix
- The Opportunity Cost Matrix
- Lift calculation
- The impact/complexity analysis matrix
- Value chain analysis

Demystifying Machine Learning

- Fundamentals of ML algorithms
- Learning to diagnose and evaluate the work of a data scientist
- Algorithmic exploration, an exploration from within, without mathematical formula:
 - Penalized linear regression
 - From trees to forests, boosting
 - Neural networks and deep learning
- Decoding the work of a data scientist and managing a data project
- Big data ecosystem

Introduction to Data Visualization

- Data visualization concepts
 - Tufte rules
- Example presentation
 - History
 - Modern
- Technology listing
 - Web
 - Programmative

Big Data technical environment

- Notion of networks
 - Routers, ports, servers, ssh
- Introduction to the Hadoop ecosystem
- Spark emerges as Hadoop's replacement
- Overview of modern infrared components
- Example of a website architecture, and log collection as Kafka

Data project management

- Common pitfalls
- "Back to the feature"
- Produce results that the business can understand
- Scrum management

GDPR (RGPD)

- History of CNIL
 - Regulations
 - The evolution of standards
 - Market modification
- Data rights abroad
 - America VS Europe
 - Harmonizing European policies
- the RGPD
 - Issues
 - Application
 - Sanctions

Companies concerned

This training course is aimed at both individuals and companies, large or small, wishing to train their teams in a new advanced computer technology, or to acquire specific business knowledge or modern methods.

Positioning on entry to training

Positioning at the start of training complies with Qualiopi quality criteria. As soon as registration is finalized, the learner receives a self-assessment questionnaire which enables us to assess his or her estimated level of proficiency in different types of technology, as well as his or her expectations and personal objectives for the training to come, within the limits imposed by the selected format. This questionnaire also enables us to anticipate any connection or security difficulties within the company (intra-company or virtual classroom) which could be problematic for the follow-up and smooth running of the training session.

Teaching methods

Practical training: 60% Practical, 40% Theory. Training material distributed in

to all participants.

Organization

The course alternates theoretical input from the trainer, supported by examples, with brainstorming sessions and group work.

Validation

At the end of the session, a multiple-choice questionnaire verifies the correct acquisition of skills.

Sanction

A certificate will be issued to each trainee who completes the course.