# KEYRUS ACADEMY

## FULL CATALOGUE 2023



keyrus
make data matter

# TABLE OF CONTENTS

## DATA SCIENCE AND ENGINEERING TRAINING

## DATA MANAGEMENT AND DATA GOVERNANCE

# DATA VISUALIZATION WITH POWER BI

# DISCOVER Training

# SCRUM Training

# Trainings on Demand

# Contact Details

> *I enjoyed the courses being both a student and an instructor. It was one of the most interesting things I did last year! The learning curve is rewarding, and the interaction great!*
>
> Data Governance Consultant
> Brussels, 2023

# WELCOME
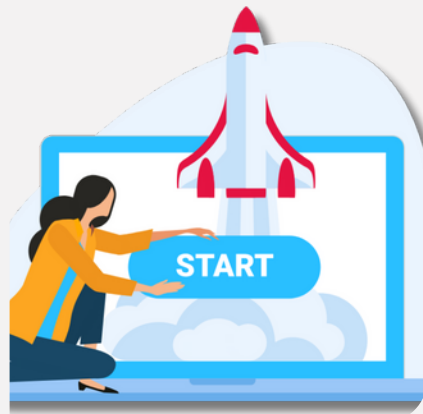## to our catalogue for Data Science and Data Engineering Trainings

As the data deluge continues to accelerate, organizations and individuals are facing new challenges. In today's fast moving data and digital landscape, figuring out how to leverage new paradigms and breakthrough technology innovations into concrete business value and competitive advantages has become a central question.

This evolution requires organizations to rethink their relationship with data, developing new strategies to leverage all relevant data sources regardless of their nature and to evolve their business models.

Here the Data Scientist and the Data Engineer play a crucial role, helping companies to bridge the gap between data, technology, advanced analytics and business issues. However, despite a rapid increase in the maturity and awareness around those topics across all industries, finding and training Data Scientists often remains a showstopper.

Therefore, in order to accelerate the transition of high-skilled academic profiles to the business and to allow companies to reap the benefits of the Big Data wave.

Keyrus has developed a unique program combining IT, technology, applied mathematics, and business context awareness build up to train young graduates or talented professionals with a first experience in data mining, analytics or data science.

START

# OVERVIEW TRAININGS

## Data Science Track

Bring your data science team to the next level by taking on our Data Science Training Track. Learn the necessary tools to undertake your data science projects, implement machine learning algorithms and models to improve decision-making.

## Data Engineering Track

This track is aimed at teams in charge of the design, implementation, and maintenance of their company's data analytics infrastructures. Learn about state-of-the-art technologies and tools to ingest and process your data.

## Big Data Track

Bring your data science team to the next level by taking on our Data Science Training Track. Learn the necessary tools to undertake your data science projects, implement machine learning algorithms and models to improve decision-making.

50% Practice    50% Theory

# **O**RGANISATION

## D.S.T.

### DATA SCIENCE TRACK

> *Innovation is about bringing value to your business, not getting lost in a vast jungle of methods and algorithms.*



**keyrus**
make data matter

### HOW TO SUBSCRIBE

Email us at be.bootcamp@keyrus.com
Or call us at 02 706 03 00
Subscription to individual courses and modules is possible

### PRICING

You will find pricing for each individual course listed in below catalog

### PRACTICALITIES

Bring your own laptop with Wi-Fi access
Sandwich lunch is included

### LOCATION

At Keyrus office (Nijverheidslaan 3/2 - 1853 Strombeek-Bever) OR At the Client's office (see pg. 32, trainings on demand)

### GUARANTEED SERVICE

Content can be adapted to the needs of the client

# INTRODUCTION TO UNIX & GIT

## DESCRIPTION
- Some experience with Unix server management and version control systems is a must to develop your own code and deploy cooperative work on data science initiatives. The aim of this training is to run you through some common IT tools you'll have to deal with in your data science projects.
- At the end of this course, you will be comfortable working on remote Unix servers and get the basics of Unix administration for the management and configuration of such servers.
- We will also teach you to set up and manage your own Git repositories to coordinate work with your colleagues and clients, to consistently keep track of your own code changes over time.
- Reduce churn rate and reach customers at the right time.

## DETAILED CONTENT
- Unix general overview and reminders
- Shell in syntax and commands
- Basics of Unix administration and management
- Networking
- General introduction to version control
- Key components of Git architecture
    - Git index
    - Commits
    - Branches

## PRACTICAL DETAILS

### Pricing
- 350€ for registration to one of our organized trainings at the Keyrus office (limited seats)
- 500€ for a session held at your premise (price per participant/min. 5 participants)
- All prices ex VAT

### Duration
- 1 day

### Pre-requisites
- Basic experience with the Unix terminal

# PYTHON FOR DATA SCIENCE

## DESCRIPTION
- Python is no doubt one of the most widely used programming languages in the data science industry.
- Through a combination of theory and hands-on exercises, we intend to provide you with a pragmatic experience with Python in data-related projects.
- In this training, you will gain the skills needed to get started with Python and launch your data science projects.

## DETAILED CONTENT
- Introduction to the language
- Data types
- Conditional structures
- Coding conventions
- Numpy
- Pandas
- Object Oriented Programming
- Visualization
- Regex
- Multiprocessing
- Web scrapping

## PRACTICAL DETAILS

### Pricing
- 1.150€ for registration to one of our organized trainings at the Keyrus office (limited seats)
- 1.500€ for a session held at your premise (price per participant/min. 5 participants)
- All prices ex VAT

### Duration
- 3 days

### Pre-requisites
- Basic experience with the Unix terminal and programming are recommended

# ADVANCED PYTHON FOR DATA SCIENCE

**DESCRIPTION**
- Python is no doubt one of the most widely used programming languages in the data science industry.
- Through a combination of theory and hands-on exercises, we intend to provide you with a pragmatic experience with Python in data-related projects.
- In this training, you will gain the skills needed to bring your data science Python skills to the next level.

**DETAILED CONTENT**
- Data types
- Packages
- String manipulation
- Predictive setup
- Scikit-learn
- Scoring
- Practical case

**PRACTICAL DETAILS**

**Pricing**
- 350€ for registration to one of our organized trainings at the Keyrus office (limited seats)
- 500€ for a session held at your premise (price per participant/min. 5 participants)
- All prices ex VAT

**Duration**
- 1 day

# R PROGRAMMING LANGUAGE

**DESCRIPTION**
- When talking about data science, it is often said that the R programming language is second only to Python.
- However, when talking about statistical analysis, time series, and its fantastic tools to communicate results, R becomes the tool of choice to integrate into your data scientist background.
- Through a combination of theory and hands-on exercise, we intend to provide you with a pragmatic and hands-on experience with the R programming language.
- In this training, you will gain the skills needed to get started on your data-related projects with R.

**DETAILED CONTENT**
- Introduction to the language
- Data analysis with the tidyverse
- Data Ingestion
- Data visualization
- R packages

**PRACTICAL DETAILS**

**Pricing**
- 700€ for registration to one of our organized trainings at the Keyrus office (limited seats)
- 1.00€ for a session held at your premise (price per participant/min. 5 participants)
- All prices ex VAT

**Duration**
- 2 days

**Pre-requisites**
- Basic experience with the Unix terminal and programming are recommended

# DATA VISUALISATION WITH PYTHON

## DESCRIPTION
- Needless to say, good data visualization skills are not only needed for reporting but also when building and monitoring your models.
- In the following courses, we run you through the most successful packages for data visualization with Python and R.
- As usual, with hands-on exercises and examples to make you started on your own data projects.

## DETAILED CONTENT
- General introduction to data visualization
  - Static visualization
  - Matplotlip
- Pandas
  - Seaborn
  - Interactive visualization
- Plotly, Bokeh, Ipywidgets
- Dynamic graphs and webapps

## PRACTICAL DETAILS

**Pricing**
- 400€ for registration to one of our organized trainings at the Keyrus office (limited seats)
- 550€ for a session held at your premise (price per participant/min. 5 participants)
- All prices ex VAT

**Duration**
- 1 day

**Pre-requisites**
- Programming with Python

# DATA VISUALISATION WITH R

## DESCRIPTION
- When talking about data science, it is often said that the R programming language is second only to Python.
- However, when talking about statistical analysis, time series, and its fantastic tools to communicate results, R becomes the tool of choice to integrate into your data scientist background.
- Through a combination of theory and hands-on exercise, we intend to provide you with a pragmatic and hands-on experience with the R programming language. In this training, you will gain the skills needed to get started on your data-related projects with R.

## DETAILED CONTENT
- Visualization with R with standard tools
- Advanced visualization with ggplot2
- A sneak peek of what else is possible
- Introduction to Rshiny

## PRACTICAL DETAILS

Pricing
- 400€ for registration to one of our organized trainings at the Keyrus office (limited seats)
- 550€ for a session held at your premise (price per participant/min. 5 participants)
- All prices ex VAT

**Duration**
- 1 day

**Pre-requisites**
- Programming with R

# MACHINE LEARNING

**DESCRIPTION**
- In recent years, Machine Learning (ML), as the right hand of Artificial Intelligence, has surged in all sectors of our society. This training will bring you on a journey through the vast jungle of machine learning algorithms. It will give you the necessary tools to understand the underpinnings of a large series of algorithms as well as provide you the right skills to code and calibrate the models as well as compute and understand their performance.
- This unique ML class will mix theoretical knowledge with business-related experience. As such, it has the pretension to go beyond what any academic class, online learning, or book could ever cover at once. Every theoretical topic will be illustrated with examples and further thoughts using our business expertise.
- The class will eventually finish with a Kaggle-like competition where you will be able to demonstrate your true ML skills!

**DETAILED CONTENT**
- Introduction to Machine Learning
- Machine Learning Project Workflow
- Regularization
- Supervised Machine Learning algorithms
- Unsupervised Machine Learning algorithms
- Hands-on: in class ML competition

**PRACTICAL DETAILS**

**Pricing**
- 400€ for registration to one of our organized trainings at the Keyrus office (limited seats)
- 550€ for a session held at your premise (price per participant/min. 5 participants)
- All prices ex VAT

Duration
- 1 day
Pre-requisites
- Programming with Python

# TEXT MINING

**DESCRIPTION**
- In a society where social networks, online customer reviews, and blog articles are thriving, Text Mining techniques had to reach new levels of comprehension and depth to exploit this rich and complex information.
- In this training, you will get a broad overview of the existing techniques and learn how to use them through hands-on exercises.
- Upon completion of this training, you will have a clear insight on the different existing techniques and their relevancy for your data environment and applications.
- You will also gain practical experience in the implementation of a text mining workflow.

**DETAILED CONTENT**
- Tokenizing text and Part-of-Speech tagging
- Preprocessing: stemming and lemmatization
- Natural Language Processing: from raw text data to structure objects
- Word2Vec
- Topic Modelling
- Latent Dirichlet Allocation
- Clustering
- Text Classification and Sentiment Analysis

**PRACTICAL DETAILS**

**Pricing**
- 400€ for registration to one of our organized trainings at the Keyrus office (limited seats)
- 550€ for a session held at your premise (price per participant/min. 5 participants)
- All prices ex VAT

**Duration**
- 1 day

**Pre-requisites**
- Basic experience with the Unix terminal and programming are recommended

# FORECASTING

**DESCRIPTION**
- Forecasting is a crucial component in various fields of concern, no matter if you are starting a new business or if you are part of a large organization.
- The correctness of management decisions, estimation of financial requirements, and plan formulation heavily rely on accurate forecasting of either sales or financial forecasting.
- In this training, we assess and present the latest and standard methods you can use to make proper forecasts adapted to your needs, and help you face these issues. Upon completion of this training, you will have a clear insight into the different existing techniques and their relevancy for your forecasting use case.

**DETAILED CONTENT**
- Time series visualization and data preparation
- Forecasting on time series data

**PRACTICAL DETAILS**

**Pricing**
- 350€ for registration to one of our organized trainings at the Keyrus office (limited seats)
- 500€ for a session held at your premise (price per participant/min. 5 participants)
- All prices ex VAT

**Duration**
- 1 day

**Pre-requisites**
- Programming with R

# DEEP LEARNING

**DESCRIPTION**
- While Artificial Intelligence is gradually transforming various aspects of our lives, many ignore that one particular technique is at the core of the spectacular changes we have witnessed in recent years.
- This technique is called 'Deep Learning', a type of machine learning that uses algorithms based on the way the human brain operates (artificial neural networks) to solve complex tasks.
- This course gives you a complete understanding of the mechanisms behind neural networks, by receiving a sound description of the main theoretical concepts and by performing hands-on coding exercises. We will use the most recent technologies to build our own algorithms and use them to solve common business problems.

**DETAILED CONTENT**
- Introduction to Deep Learning and its use cases
- Theoretical foundations of Deep Learning
- Overview of Neural Networks features
- Implementation of your own Neural Network with Keras and Tensorflow
- Hands-on: Computer Vision
- Hands-on: Natural Language Processing

**PRACTICAL DETAILS**

**Pricing**
- 350€ for registration to one of our organized trainings at the Keyrus office (limited seats)
- 500€ for a session held at your premise (price per participant/min. 5 participants)
- All prices ex VAT

**Duration**
- 1 day

**Pre-requisites**
- Programming with R

# D.E.T.

## DATA ENGINEERING TRACK

" A data scientist is only as good as the data they have access to: build the right playground for your data science teams with our data engineering track. "

# UNIX

## DESCRIPTION
- Since its creation in the 1970s, Unix has become one of the operating systems of choice for business servers, be them on-premise or in the cloud.
- In particular, as a data scientist/engineer, most projects require at least basic interaction with Unix servers, ranging from simple file management to server administration in Big Data clusters.
- Follow this Unix training and gain the skills needed to administer and work remotely on your Unix servers.

## DETAILED CONTENT
- Unix general overview and reminders
- Shell in depth syntax and commands
- Exploring Administrative Commands, Configuration Files, and Log Files
- Management (software, user accounts, disks and filesystems)
- Server Administration
- Network Administration
- Configuration of some example services (web server, ftp, ..)
- Linux Security

## PRACTICAL DETAILS

**Pricing**
- 800€ for registration to one of our organized trainings at the Keyrus office (limited seats)
- 1.100€ for a session held at your premise (price per participant/min. 5 participants)
- All prices ex VAT

**Duration**
- 1 day

**Pre-requisites**
- Basic knowledge of the Unix shell is recommended

# GIT

## DESCRIPTION
- Version control systems are a must to keep track of history, coordinate and integrate cooperative software development. Among available technologies, Git is one of the most popular distributed version control systems to coordinate those tasks.
- At the end of this training, you will be able to set up and manage your own Git repositories, either local or remote. A combination of theoretical concepts and exercises will show you how to coordinate work on source code with your co-workers by consistently keeping track of your own and everybody's code changes over time.

## DETAILED CONTENT
- Introduction to Version Control Systems
- Key components of Git architecture
- Git index
- Commits
- Branches
- Remote repositories
- Typical git life cycle
- GUI

## PRACTICAL DETAILS

**Pricing**
- 350€ for registration to one of our organized
   trainings at the Keyrus office (limited seats)
- 500€ for a session held at your premise (price per participant/min. 5 participants)
- All prices ex VAT

**Duration**
1 day
**Pre-requisites**
Basic knowledge of the Unix shell

# PYTHON FOR DATA ENGINEERS

## DESCRIPTION
- Python has become the language of choice within the data industry, and in particular within the context of production and deployment of your data models and pipelines.
- Through a combination of theory and hands-on exercise, this training will provide you with a pragmatic and hands-on experience of Python with particular emphasis on data processing and transformation.

## DETAILED CONTENT
- Introduction to the language
- Data processing with Python
- Data processing with Pandas
- (Big) data processing and optimization
- Alternatives to Pandas
- Data visualization and monitoring

## PRACTICAL DETAILS

**Pricing**
- 700€ for registration to one of our organized trainings at the Keyrus office (limited seats)
- 1.000€ for a session held at your premise (price per participant/min. 5 participants)
- All prices ex VAT

**Duration**
1 day
**Pre-requisites**
Basic knowledge in programming and data preparation/management. Basic knowledge of Python is advised.

# INTRODUCTION TO DATABASES

**DESCRIPTION**
- Databases underlie technology used in most industries from telecommunication and banking to any software maintaining persistent information.
- The so-called relational model has been widely used in the last decades throughout the industry. While still prevalent, the advent of the internet and big data technologies has seen the emergence of a NoSQL database to handle and manage semi-structured and unstructured data.
- This training gives you a general introduction to the world of databases and their concepts, covering both the traditional relational model and some widely used NoSQL database management systems. These concepts are illustrated with hands-on practice on SQL servers and the document-oriented database MongoDB.

**DETAILED CONTENT**
- General introduction to databases and SQL hands-on
- Database models
- Data warehouse models
- Query languages
- Hands-on with SQL server

- NoSQL with MongoDB
- Basics and setting up
- MongoDB operations, aggregation framework
- MongoDB and data science and applications

**PRACTICAL DETAILS**

**Pricing**
- 700€ for registration to one of our organized trainings at the Keyrus office (limited seats)
- 1.000€ for a session held at your premise (price per participant/min. 5 participants)
- All prices ex VAT

**Duration**
- 1 day

**Pre-requisites**
- Basic knowledge in programming and data preparation / management
- Basic knowledge of Python is advised

# INTRO DOCKER & KUBERNETES

**DESCRIPTION**
- Containerization has dramatically changed the way we create and deploy applications. Docker has become the de facto standard platform-as-a-service to deploy and run your applications in containers while Kubernetes allows you to easily orchestrate and run many containers in coordination.
- On the first day of this training, we will give you a general introduction to virtualization and container technology, as well as hands-on experience with Docker. On the second day, we will start working on Kubernetes and teach you how to use some of the functions this ecosystem can bring you.

**DETAILED CONTENT**

**Docker**
- General introduction
- Dockerfile
- Docker command
- Docker compose
- Orchestration

**Kubernetes**
- General introduction
- Dockerfile
- Docker command
- Docker compose
- Orchestration

**PRACTICAL DETAILS**

**Pricing**
- 700€ for registration to one of our organized trainings at the Keyrus office (limited seats)
- 1.000€ for a session held at your premise (price per participant/min. 5 participants)
- All prices ex VAT

**Duration**
- 2 day

**Pre-requisites**
- Basic knowledge of the Unix terminal

# INTRODUCTION TO CLOUD SERVICES

**DESCRIPTION**
- The Cloud is transforming the way organizations use and manage technology, deeply impacting business processes, governance, and IT systems.
- This training is a first introduction to cloud services, with emphasis on two of the major players on the market: Microsoft Azure and Amazon Web Services.
- By following this training, you will get a good understanding of the working architecture with hands-on practice on the standard services offered by these providers. The aim is also to make you feel conformable and enable you to build on this course to use the services that suit your needs.

**DETAILED CONTENT**
**Microsoft Azure**
- Introduction to Azure
- Azure architecture and service guarantees
- Azure accounts
- Managing services with the Azure portal
- Azure compute options• Azure networking

**Kubernetes**
- AWS main services
- Security and compliance
- Pricing
- Well-Architected framework

**PRACTICAL DETAILS**

**Pricing**
- 700€ for registration to one of our organized
  trainings at the Keyrus office
(limited seats)
- 1.000€ for a session held at your premise (price per participant/min. 5 participants)
- All prices ex VAT

**Duration**
- 2 day

**Pre-requisites**
- Basic knowledge internet technologies and Unix A background in programming is recommended

# B.D.T.

## BIG DATA TRACK

> " *A data scientist is only as good as the data they have access to: build the right playground for your data science teams with our data engineering track.* "

# DATA ARCHITECTURE – FOUNDATION

**DESCRIPTION**
- Data architecture is often perceived as vague or focused on drawing data models. In fact, there's much more to this practice. The data architect is a central role in the data domain because data architecture is involved in every other domain with data management.
- This course will give you the toolset from the ground up, to make informed decisions when working on data initiatives. A solid foundation will help the enterprise and the projects fulfill the business needs. This course offers you this solid foundation.
- By zooming in and zooming out of specific domains, relating to architecture, you will have a holistic perspective on the data domain.

**DETAILED CONTENT**
- Introduction
- Fundamental Data Theory
- Data Modeling
- Data Integration
- Data Consumption
- Data Architecture
- Deployment & Change Management
  Security Basics

**PRACTICAL DETAILS**

**Pricing**
- 2.000€ for registration to one of our organized trainings at the Keyrusoffice (limited seats)
- 2.250€ for a session held at your premise (price per participant/min. 5 participants)
- All prices ex VAT

**Duration**
- Estimated 4-5 Days

**Pre-requisites**
- None, but IT background and basic data experience are useful in this course

# PYSPARK

**DESCRIPTION**
- Pyspark is the standard tool for big data processing, in memory and in parallel on your Hadoop cluster. In this training, we present the working concepts of Spark and give a detailed presentation of the data processing and analysis methods available within the PySpark language context.
- This training is aimed at giving you the ability to independently launch your big data analytics projects with PySpark:
  - By building on the exercises and examples which are part of the training
  - By giving you the relevant information to look out for solutions to your big data analytics requirements

**DETAILED CONTENT**
- Why Spark?
- Basics of PySpark
- Spark operations
- Spark API
- Applications
- Spark SQL and Dataframes
- Other libraries

**PRACTICAL DETAILS**

**Pricing**
- 800€ for registration to one of our organized trainings at the Keyrus office (limited seats)
- 1.100€ for a session held at your premise (price per participant/min. 5 participants)
- All prices ex VAT

**Duration**
- 2 days

**Pre-requisites**
- Programming with Python

# HBASE

## DESCRIPTION
- This training will provide you with a thorough introduction HBase, a high-performance, distributed column-oriented data store built for Apache Hadoop, providing real-time read/write access to large datasets.
- HBase allows for storing a massive amount of data in the form of a table consisting of billions of rows and millions of columns.
- This HBase training is specifically designed for Big Data developers, with the right combination of theoretical presentation and hands-on exercises.

## DETAILED CONTENT
- General Introduction to HBase
- HBase Data Model
- Basic Usage in the Shell
- Schema Design: Column-oriented Data Store
- HBase Architecture
- Advanced Usage: APIs and Connectors
- Performance Tuning

## PRACTICAL DETAILS

### Pricing
- 800€ for registration to one of our organized trainings at the Keyrus office (limited seats)
- 1.100€ for a session held at your premise (price per participant/min. 5 participants)
- All prices ex VAT

### Duration
- 2 days

### Pre-requisites
- Basic knowledge of the Hadoop ecosystem

# KAFKA

## DESCRIPTION
- Apache Kafka is one of the most powerful and widely used distributed streaming-processing platforms. Fault-tolerant and highly scalable, it is the tool of choice for high-throughput, low-latency handling of real-time data feeds, such as event or commit logs.
- This training will give you a comprehensive knowledge of Kafka's architecture as well as hands-on experience with the platform. You will learn to get the most out of this technology as:
  - A distributed messaging system
  - A storage component for streamed data
  - An integration system for your Big Data applications

## DETAILED CONTENT
- Introduction to Apache Kafka
- Kafka Architecture
- Data Flow
- Kafka CLI
- Kafka Extended APIs
- Kafka in Real-World
- Kafka Case studies

## PRACTICAL DETAILS

### Pricing
- 800€ for registration to one of our organized trainings at the Keyrus office (limited seats)
- 1.100€ for a session held at your premise (price per participant/min. 5 participants)
- All prices ex VAT

### Duration
- 2 days

### Pre-requisites
- Basic knowledge of the Hadoop ecosystem

# SPARK STREAMING

## DESCRIPTION
- The Spark Streaming module extends the Spark batch infrastructure to deal with Big Data for real-time analysis such as trends and breaking news. Lying on top of the Hadoop ecosystem, it runs your jobs in parallel on a cluster ensuring fault-tolerance and scalability.
- In this training, we will cover the basic concepts of streaming data processing using the same basic Spark transformations that work with batch processing. You will be given hands-on exercises to illustrate those concepts and to feel comfortable when starting your own streaming data processing projects.

## DETAILED CONTENT
- Introduction to Data Streaming
- Spark Streaming Architecture

Implementation:
- Sources and Sinks
- Event Time vs Processing
- Sorting and Dedup
- Stateful Processing
- Transformations
- SQL on Streams

## PRACTICAL DETAILS

### Pricing
- 800€ for registration to one of our organized trainings at the Keyrus office (limited seats)
- 1.100€ for a session held at your premise (price per participant/min. 5 participants)
- All prices ex VAT

### Duration
- 2 days

### Pre-requisites
- Basic knowledge of the Hadoop ecosystem and PySpark

# DESIGNING OOZIE WORKFLOWS

## DESCRIPTION
- Learn how to use Oozie to manage and orchestrate jobs running simultaneously on your Hadoop cluster by following this training. The Oozie workflow scheduler integrates with all the tools of the Hadoop ecosystem – MapReduce, Hive, Pig, and Sqoop – but also supports system-specific jobs such as Java coding programs and shell scripts.
- This training covers Oozie's architecture and will teach you how to write your own workflows to build your Big Data pipelines.

## DETAILED CONTENT
- Core Concepts
- Designing Oozie Workflows
- Oozie Actions
  - Java
  - Shell
  - Sqoop
  - Hive
- Designing Coordinators

## PRACTICAL DETAILS

### Pricing
- 400€ for registration to one of our organized trainings at the Keyrus office (limited seats)
- 550€ for a session held at your premise (price per participant/min. 5 participants)
- All prices ex VAT

### Duration
- 2 days

### Pre-requisites
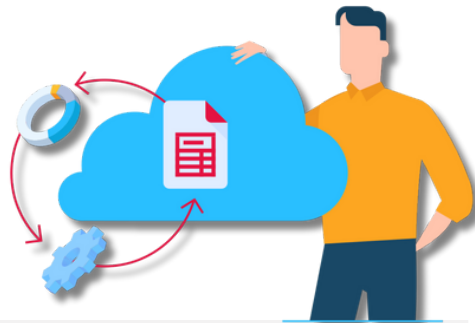- Basic knowledge of the Hadoop ecosystem

# DEVOPS TRACK-AUTOMATION & CI/CD

**DESCRIPTION**

- Learn how you and your company can benefit from state-of-the-art DevOps tools and methodology and move towards faster releases of their software and more reliable processes.
- In this training, we will introduce you to the main DevOps concepts for automation and continuous integration and delivery (CI/CD) and will also provide you with an overview of the major tools to implement those concepts.
- At the end of this training, you will be able to set up your automated code publication and integration system, in a functional production environment.

**DETAILED CONTENT**

- General Introduction to DevOps
- Infrastructure as Code
- Introduction to Ansible
- General Overview of CI/CD
- Jenkins

**PRACTICAL DETAILS**

**Pricing**

- 800€ for registration to one of our organized trainings at the Keyrus office (limited seats)
- 1.100€ for a session held at your premise (price per participant/min. 5 participants)
- All prices ex VAT

**Duration**
2 days

**Pre-requisites**
Basic programming knowledge
Basic knowledge of version control

> *Learned so much, finally understand the difference between data governance and data management.*
>
> Data Governance Consultant
> Brussels, 2023

# WELCOME TO OUR TRACK:
## "Data Management & Data Governance"

Managing data as a vital enterprise asset. This is the aim of many organizations, in order to capture insights about their customers, products, and services so they can continue to innovate, to grow their business, and become more relevant and personalized in their approach. However, in spite of this recognition, few organizations succeed in actively managing their data as an asset from which they can derive ongoing value. Some have difficulties on how to organize for Data Management, others on which areas to cover and where to start.

In this learning track, we will kick-start your Data Management journey by focusing on how applying best practices in the domains of Data Governance, Analytics Governance, Metadata Management, Data Quality Management, and Master Data Management can support your organization to start managing data as an asset.

After completion of the different modules, you will understand how to organize for Data Management activities, what supporting tools/artefacts can help you in this journey, and how these activities will ultimately contribute to the realization of your organization's goals.

This track is aimed at both technical and non-technical people already active in the data domain or with a specific affinity for data.

**50% Practice**
**50% Theory**

# OVERVIEW OF THE PROPOSED MODULES

This training pack contains 5 modules, each of them corresponding to the expertise domains covered by Keyrus.

These 5 modules can be given together in a package of 5 days, or more "à la carte" according to your domain of interest.

**1 DATA GOVERNANCE**
Focuses on how decisions are made about data and how people and processes are expected to interact in relation to data.

**2 ANALYTICS GOVERNANCE**
Describes how to organize the analytical activities (BI/Data Science/self-service) in the most efficient way in order to generate the largest impact possible for the entire oragnization.

**3 METADATA MANAGEMENT**
How to manage the "Who, What, Why, When & How" of your data, and make this information transparent to the broader organization.

**4 ANALYTICS GOVERNANCE**
Describes how to organize the analytical activities (BI/Data Science/self-service) in the most efficient way in order to generate the largest impact possible for the entire oragnization.

**5 MASTER DATA MANAGEMENT**
How to ensure the uniformity, accuracy stewardship, semantic consistency and accountability of the enterprise's official shared master data assets..

**6 CUSTOMIZED TRAINING**
We can also offer more tailored trainings for "on demand" specific needs.

**keyrus**
make data matter

# MODULES

## MODULE # 01

Data Governance & Data Management

## MODULE # 02

Analytics Governance

## MODULE # 03

Metadata Management

## MODULE # 04

Data Quality Management

## MODULE # 05

Master Data Management

## MODULE # 06

Data Literacy

**keyrus**

---

## DATA GOVERNANCE

Focuses on how decisions are made about data and how people and processes are expected to interact in relation to data.

## MODULE 1

## DATA GOVERNANCE & DATA MANAGEMENT

### DESCRIPTION

Nowadays, companies must live in a data-driven world where an increased volume of many different kinds of data is coming into our systems at a speed never seen before. This massive amount of data leads to opacity, inefficiency, and lack of transparency.

Therefore, to solve these issues data needs to be managed and governed.

All organizations already govern their data to some degree, regardless of whether they have a formal Data Governance function. Research shows that organizations that formalize their Data Governance are in a better position to rapidly increase the value they get from their data assets and to sustain this value generation in time. This is because Data Governance focuses on how decisions are made about data and how people and processes are expected to interact in relation to data. It enables an organization to be more efficient in dealing with its data, accelerate its data exploitation, and unlock new business value.

This session will start with some basics about why Data Management should be a cornerstone of your company's strategy, how it differs from Data Governance, and why you will need both to achieve your company's strategic objectives.

Participants will then gain a better understanding of the essential concepts of Data Governance, and how an organization can benefit from formalizing and improving it. We will explain the most important Data Governance frameworks and how to practically organize for and around data in an organization. The needed roles and linked responsibilities will be explained, as well as typical Data Governance processes to support a formalized way of working.

Concrete examples will support the concepts and frameworks shown.

## DETAILED CONTENT

- Demystifying the concepts of Data Governance & Data Management
- Why embarking on a Data Governance & Data Management journey
- 10 important considerations before getting started
- Defining Data Governance more in detail
- Essential concepts of Data Governance
- Data Governance Frameworks
    - Operating Models
    - Data Governance principles
    - Typical governance bodies
    - Typical roles & responsibilities
    - Data Governance processes
- Non-invasive Data Governance
- How to get started with Data
- Governance Key success factors

## PRACTICAL DETAILS

**Duration**
1 day
Pre-requisites
None, however, basic data experience is useful, in this co-race

# ANALYTICS GOVERNANCE

Describes how to organize the analytical activities (BI/Data Science/self-service) in the most efficient way in order to generate the largest impact possible for the entire organization.

# MODULE 2

## DATA GOVERNANCE & DATA MANAGEMENT

### DESCRIPTION

Organizing internally to manage data on a daily basis is a challenge, organizing efficiently on how to best exploit data is yet another one. In a context in which an increasing number of companies aim to accelerate and increase the exploitation of their data (e.g. through reporting, descriptive or predictive analytics), the demand for talent with the appropriate data analytics skills is soaring. Yet it remains a scarce resource on the market, which forces companies to use their analytical talent and competencies in the most effective way in order to generate the largest impact possible for the entire organization.

This module will provide insights on how to organize internally to ensure the right analytical skills and competencies are involved in the right type of (analytical) activities for the right scope. We will introduce some frameworks on how to organize Business Intelligence activities, Data Science activities, and self-service activities. We will address the required processes to support these frameworks, as well as the needed roles and linked responsibilities.

We will also address how these frameworks can be embedded into the larger Data Governance within an organization. in which an increasing number of companies aim to accelerate and increase the exploitation of their data (e.g. through reporting, descriptive or predictive analytics), the demand for talent with the appropriate data analytics skills are soaring. Yet it remains a scarce resource on the market, which forces companies to use their analytical talent and competencies in the most effective way in order to generate the largest impact possible for the entire organization.

This module will provide insights on how to organize internally to ensure the right analytical skills and competencies are involved in the right type of (analytical) activities for the right scope. We will introduce some frameworks on how to organize Business Intelligence activities, Data Science activities and self-service activities. We will address the required processes to support these frameworks, as well as the needed roles and linked responsibilities. We will also address how these frameworks can be embedded into the larger Data Governance within an organization.

## DETAILED CONTENT

- Analytics Governance definition & difference with Data Governance
- Business drivers and challenges behind Analytics Governance
- Analytics Governance Frameworks, including roles and responsibilities
  - Analytics Governance
  - Data Science Approach
- Analytics Governance Processes
  - High-level description of processes
  - Detailed process & instructions definition
- Implementing Analytics Governance within an organization
- Analytics Governance and Technology
- Key success factors
- Use-case example

## PRACTICAL DETAILS

**Duration**
- 1 day

**Pre-requisites**
- Introduction to Data Governance & Data Management Module

keyrus
make data matter

## METADATA MANAGEMENT

How to manage the "Who, What, Why, When & How" of your data, and make this information transparent to the broader organization.

## MODULE 3

## DATA GOVERNANCE & DATA MANAGEMENT

### DESCRIPTION

Many companies struggle to align internally on unique definitions for business terms (e.g. "What is the definition of a customer?"). For others, it is a challenge to have clear KPI definitions. Data Scientists and Reporting Officers can find it hard to gain a view of the available data assets, and what this data is exactly about. When documentation is available, it is often a challenge to keep data documentation up to date. Sounds familiar? In order to manage data effectively within an organisation, it is important to have a view on and to understand its metadata (i.e. the information about the data, such as its data structure, a data definition, how data flows between systems, who is the owner of the data, where the data is located...). Well-managed metadata has the potential to accelerate dramatically the exploitation of your data assets, as your data users will no longer have to look endlessly to find the right data for the right purpose. This module will provide insights on what are the typical metadata management artefacts (e.g. Business Glossary and Data Dictionary) that can support your organization's metadata management approach. It is, however, important to organize around these artefacts to extract value from them, this is why we will also introduce a metadata governance framework and explain which roles and responsibilities should typically be foreseen. As the current offer for appropriate tooling on the market is soaring, we will also provide guidance on how to navigate through it with a quick market review and by explaining the most essential functionalities needed to support your metadata approach.

## DETAILED CONTENT

- Metadata definition & concepts
- Types of Metadata
- Metadata management artefacts
  - Business Glossary
  - Data Dictionary
  - Data Catalog
  - Data Lineage & impact analysis
- Linking Metadata & Data Governance
- Metamodels
- The tooling aspect
  - Typical functionalities
  - Market scan
  - Demo
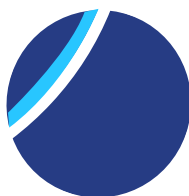- Implementing Metadata Management within an organisation

## PRACTICAL DETAILS

**Duration**
- 1 day

**Pre-requisites**
- Introduction to Data Governance & Data Management Module

# DATA QUALITY MANAGEMENT

How to make your data fit for purpose and trustworthy, combining the right people, processes, and tools.

## DATA GOVERNANCE & DATA MANAGEMENT

## DESCRIPTION

Extracting data and exploiting it to unlock additional business value has an underlying assumption that the data itself is reliable and trustworthy. This does not mean that data should be perfect, but rather that the data should be fit for its intended purpose. Even so, as no organization has perfect business or technical processes, nor perfect data management practices, all organizations experience problems related to the quality of their data. In this context, organizations that formally manage the quality of their data have a head start and fewer data problems compared to others.

During this session, participants will learn what Data Quality Management is about, why it is an essential part to realize your organization's strategic objectives, and how to organize for it. We will explain which roles and linked responsibilities will typically be involved, and what supporting processes should ideally be foreseen. Finally, we will also provide insights regarding available tooling on the market to support your Data Quality activities within your organization. The trainee will perform some practical exercises during this session.

## DETAILED CONTENT

- Data Quality/ Data quality Management definition
- Business drivers behind Data Quality Management
- The guiding principles for improving data quality
- Data quality dimensions
- Implementing Data Quality Management within an organisation
- Data Quality Framework, including typical roles & responsibilities
- Linking Data Quality & Data Governance
- Common Data Quality Management tools, quick market review and offered functionalities
- Use-case example & tooling demo

## PRACTICAL DETAILS

**Duration**
- 1 day

**Pre-requisites**
- Introduction to Data Governance & Data Management Module

# MASTER DATA MANAGEMENT

How to manage the "Who, What, Why, When & How" of your data, and make this information transparent to the broader organization.

## MODULE 5

# DATA GOVERNANCE & DATA MANAGEMENT

## DESCRIPTION

In this time of digitalization, data are everywhere and often siloed in many different systems, coming from many different sources and following different standards. Many companies are willing to move forward to a customer 360 view, to a customer zone, or to a better customer experience. But this is only possible when you can reconciliate and consolidate your data from different systems into one central application, through deduplication processes and Golden Record creation. This is the goal of Master Data Management. During this session, the trainee will be provided with a practical overview of what Master Data Management is all about. The theoretical aspects will be combined with real-case examples and exercises. Every topic will focus on a specific area of Master Data Management, which should provide a complete view at the end of the training

## DETAILED CONTENT

- MDM: The basics
- Software solutions
- The Master Data Model
- MDM Framework
    - STRUCTURE: Store Master Data in a data model which is structured in line with the business
    - VALIDATE: Check the quality of the data to block incorrect data or to trigger corrective actions
    - DEDUPLICATE: Detect duplicated records across different systems and use the best value from it
    - ENRICH: Improve the available data by applying rules, use of web services or data stewards

- Linking MDM & Data Governance, including typical roles & responsibilities
- Reporting and dashboarding
- Use-case example
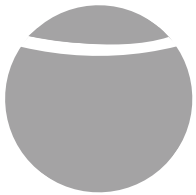
## PRACTICAL DETAILS

**Duration**
- 1 day

**Pre-requisites**
- Introduction to Data Governance & Data Management Module

# DATA LITERACY

The ability to read, write and communicate data in context, including an understanding of data sources and constructs, analytical methods and techniques applied, and the ability to describe the use case, application and resulting value.

## MODULE 6

## DATA LITERACY

## DESCRIPTION

Keyrus is now partnering with Data Literacy Academy to offer you the most comprehensive data literacy approach. But what is it exactly? According to Gartner, it is "The ability to read, write and communicate data in context." At the Data Literacy Academy, we also believe that in the end, it is about making business-focused people more data-minded, while making data-focused people more business-minded. Data literacy is at the of every successful data programme, so are you ready to make your workforce data Savvy?

Data literacy academy offers you different kinds of learning certification paths, based on the three main aspects of Data literacy: The foundation of data: Gartner says that up to 82% of employees in your departments don't know where to start with data. These courses teach fundamental data knowledge and strategy so that every department can deliver effectively with data. Work with data: There are thousands of data tools on the market, with everything from Excel to PowerBI and complex databases being used across your organisation. The Working with Data courses teach non-technical and technical team members to work effectively with data. Take action with data: Your biggest barrier to being data-driven is your employees' lack of understanding in how to convert data into action. The Taking Action with Data courses teach your team to understand, interpret, communicate, and take action effectively with data.

## DETAILED CONTENT

- Every certification path typically contains 4 modules
- Each module includes 4 sessions of 45min
- Adapted to your needs: possibility to skip/replace/customize modules and sessions
- Example of modules from the Data for Business certification track
  - Foundations of Data
  - Interpreting & communicating data
  - Tableau / Power BI training
  - Storytelling with Data
- Example of modules from the Data Specialist certification track
  - Foundations of Business
  - Scope like a Pro
  - Partnering with the Business
  - Storytelling with Data
- More details are available upon demand. More info: www.dl-academy.com

## PRACTICAL DETAILS
**Duration**
- 1 day

# D.V.

# DATA VISUALIZATION WITH POWER BI

## INTRODUCTION

Having an accurate Dashboard with relevant KPIs is essential nowadays at any management level. With Microsoft Power BI Desktop version being available for free, in-house development of such a dashboard is now accessible to anyone with a bit of Business Intelligence savviness.

## DESCRIPTION

Keyrus Power BI experts can train you on developing your own dashboard, either at a beginner level or at a more advanced level.

## DETAILED CONTENT

### MODULE # 01

Power BI Landscape, Set up of Power BI, Data sources, ETL, Query Editor

### MODULE # 02

Star schema, Create Data Model, Create First Visuals, basic rules on use of Visuals.

## MODULE # 03

Introduction to DAX, Advanced analytics, Import Visuals, Dynamic Reports, Report Optimisation

## MODULE # 04

Creating Dashboards, Paginated reports, Publish/Share Reports, Structuring workspaces, Power BI Service, Power BI App, Access Mgt, Row Level Security, Roles

### PRACTICAL DETAILS

**Duration**
- 1 day  (Module 1 and 2) or 2 days (Module 1 to 4)

**Pre-requisites**
- Using Excel

**Pricing per day**
- 700€ for registration to one of our organized trainings at the Keyrus office (limited seats)

- 1000€ for a session held at your premise (price per participant/min. 5 participants)

- All prices ex VAT

# DISCOVER Training

## TRAINING OVERVIEW

### Content:

DISCOVER is a complete toolkit for the leaders, the change agents and the professionals who want to find new ways of connecting people and empowering teams to be courageous and focused when building tomorrow's dynamic workplaces.

### Objective:

The DISC model provides a common language that people can use to better understand themselves and to adapt their behaviours with others. This can be within a work team, a sales relationship, a business development relationship, a leadership position, or any other personal relationships. As a result, you can relate and communicate even better and more effectively with other people.

### The DISCOVER training helps you and your team to:

- Increase self-knowledge: how you solve problems or achieve results, how you interact with others and show emotions, how you respond to changes or pace in the environment, how you comply with procedures and protocols, how you respond to conflict, what motivates you and what causes you stress.
- Manage more effectively customers by understanding their dispositions and priorities.
- Develop stronger sales skills by identifying and responding to customer styles.
- Facilitate better teamwork and minimize team conflict.
- Become more self-knowledgeable, well-rounded and effective leaders.

## The DISCOVER training helps you and your team to:

- Welcome & setting the scene.
- DISCOVER behavioural styles revealed.
- Personal DISCOVER report.
- Recognizing other styles.
- Adapting your style to others.
- Coaching time.

## TRAINING PRACTICALITIES

- Duration: 1 day (start at 9h00 – end at 17h00)
- Delivery date: xxx
- Number of participants per training:
  o Minimum 8
  o Maximum 12 to ensure sufficient personal attention.
- Language of training delivery: English
- A copy of the key training material and slides that will be used during the training will be sent to the training participants after the training.

## PARTICIPANT'S PREPARATION

- The training participants will be requested to fill out an on-line questionnaire before the training.
- Based on this questionnaire, the participants will receive a fully personalized profile report that will be used during the entire training.
- List of participants and their respective e-mail addresses need to be provided 3 weeks before the training date.
- The participants need to fill out on-line questionnaire to obtain personal DISC report
- Link to questionnaire will be sent minimum 1,5 week in advance (deadline for response is 1 week before training, to allow pre-reading and hand-out printing)

# SCRUM Training

## TRAINING OVERVIEW

### Content:

Today, more and more companies adopt the Agile way of working, and not only in IT departments. Among the trendy Agile frameworks, you might have heard about SCRUM. But, what is Scrum?

### Objective:

SCRUM is an agile project management framework that helps teams structure and manage their work in the most efficient way.
SCRUM comes from the world of Rugby when all the teams gets together, shoulder against shoulder, and combines all its efforts towards the same objective: the success.
SCRUM is a framework that fosters efficient collaborative work.

## The SCRUM training helps you to:

- Discover and learn about the SCRUM concepts: sprint, artifacts, ceremonies… will be clarified
- Understand how work is addressed in this framework
- Know the roles and responsibilities of Scrum Master and Product Owner
- Have a peak on the existing tools in the market
- Get certified as Scrum Master and/or Product Owner
- Manage your team work or a project more efficiently

## Training Agenda:

- Day 1 :
  - o Agile values and principles
  - o Composition of a Scrum team
  - o Scrum events part 1 / 2
- Day 1 :
  - o Scrum events part 2 / 2
  - o Artifacts
  - o Short Rehearsal
  - o Certification Exam

## TRAINING PRACTICALITIES

- Duration: 2 days (start at 9h00 – end at 17h00)
- Delivery date: xxx
- Number of participants per training:
  - o Minimum 6
  - o Maximum 10 to ensure sufficient personal attention.
- Language of training delivery: English
- A copy of the key training material and slides that will be used during the training will be sent to the training participants after the training.
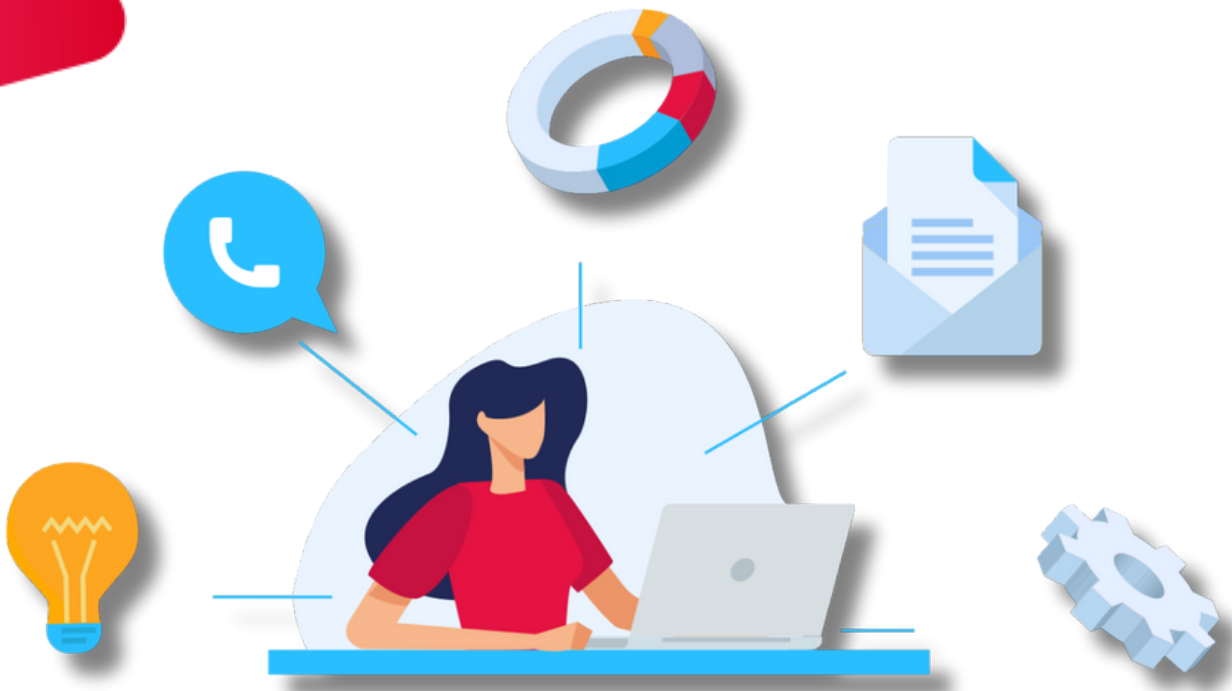
## PARTICIPANT'S PREPARATION

List of participants and their respective e-mail addresses and current role in the company need to be provided 3 weeks before the training date.

# T.D.

**keyrus**
make data matter

## Trainings on Demand?
It is possible!

"
Want us to organize a training session for you? Adapted to your demands? We are happy to do so! We create tailored and customized trainings for minimum 5 participants. Feel free to contact us in case of any questions!
"

# For inquiries, contact us:

**Address:**
Nijverheidslaan 3/2
1853 Strombeek

**Web:**
be.bootcamp@keyrus.com
www.keyrus.be

**Phone:**  +32 2 706 03 00

# THANK YOU!

keyrus

make data matter