2023

Training Guide

Machine Learning Engineering Learning Journeys

Xebia | Data

Our Promise to You

- Learn Today, Apply Tomorrow
- Authority From the Field, Like No Other
- Learn by Doing, in Hands-On Labs
- From Team to Boardroom
- Learn Your Way, at Your Pace

Go to the academy ›

Learn Today, Apply Tomorrow, and Stay in the Driver's Seat

Are the advancements in cloud and ready-to-use technologies such as Stable Diffusion and GPT-3 making data engineering and data science skills obsolete? Far from it. Without knowing what is going on under the hood, extracting maximum performance is impossible. Understanding, controlling, and utilizing the full power of these tools makes having up-to-date data and AI skills more relevant than ever.

There was a time, not so long ago, when gathering enough data to train a deep learning model — and then figure out how to train it on a budget — was an essential skill of data professionals.

Then pre-trained models and communities around them — such as HuggingFace — made it not only easy, but quick and affordable.

This begs the question: Is being proficient in data science and engineering still relevant when you can outsource the job to someone else? The answer is yes — because everything is changing again.

With all the advantages of offloading our algorithms to "someone else", it's more important than ever to stay in the driver's seat. Here are three reasons why:

(1) The lack of transparency and apparent biases in algorithmic decision-making has fuelled public outcry on numerous occasions — the Netherlands made the world news with discrimination against dual nationality families and child benefits. You can only control the algorithm if you can tweak and modify it to address the ethical issues around it. Understanding what happens under the hood, even when using AI-as-a-Service, allows you to stay in Control.

(2) In 2023, we will continue to generate TBs of data per person every month. Storing it all would be far too expensive, but we can apply machine learning to the devices that generate it and only keep the insights. More and more systems at the edge, including mobile phones, have frameworks to help you out.

(3) Transfer learning takes a highly sophisticated algorithm and adapts it to your use case and data. As it requires less training samples and computational power, it is affordable to do it in-house.

These three trends make data science and engineering skills more relevant than ever.

So, it is my pleasure to present Xebia's all-new training curriculum. Developed and taught by the very best professionals in the field with pragmatism in mind — you can apply tomorrow what you learn today. Happy learning!

Giovanni Lanzani
Managing Director Data
Xebia

Go to the academy ›

Table of Contents

Xebia Academy

Our Promise to You

At Xebia, we believe professional development goes hand-in-hand with staying happy, motivated, confident, and relevant in your job. You trust in us to help you improve your data and AI skills, and we take that very seriously.

So we strive to provide you with the best learning experience possible, by adhering to five promises:

(1) Learn Today and Apply Tomorrow
We design our programs so you can apply your newly acquired knowledge right out of the classroom. That way you can immediately increase your business value. We push the content of every training far beyond textbooks and theory. Applicability is one of our core values.

(2) Authority From the Field, Like No Other
We don't come from an abstract background; we are practitioners as much as we are teachers. All of our trainers work as consultants in the data and AI field, supporting top enterprises like Booking.com, ING, bol.com, Randstad, and Heineken. They solve challenges like yours every day, so you benefit directly from their experience.

(3) Learn by Doing, in Hands-On Labs
You learn best by doing. That's why, in every training, you can develop your skills and craft in our hands-on labs. We provide both the theory and context to get you up and running fast.
All of our courses have a 50/50 split between theory and hands-on labs.

(4) From Team to Boardroom
Built through years of working with the top enterprises in Europe, Xebia has the expertise to turn your data-driven ambition into reality. But becoming data-driven impacts your whole organization. That's why we deliver a wide range of programs suitable for all shapes and sizes—from individual teams to global workforces, as well as the boardroom.

(5) Learn Your Way at Your Pace
We deliver our curriculum through various training formats—classroom, in-company, online, or a combination. You choose the format that fits your purpose and preferred method of learning—or ask one of our academy advisors to guide you.

Xebia is an international consulting and training company, specialized in digital transformation. Xebia employs over 6,000 consultants worldwide. Xebia's values:

- People First
- Sharing Knowledge
- Customer Intimacy
- Quality Without Compromise

Xebia

Data and AI Learning Journeys

A training course should not just be a single moment of development. It works best when it's part of a broader program. At Xebia, we call these programs Learning Journeys.

Learning Journeys guide you through our curriculum and ensure the courses meet your needs and provide an effective development path. For example, our three Python courses take participants from beginner to expert, after which you can choose a specialization course to further hone their skills.

The Necessity of Learning Journeys

Without guidance, it can be challenging to invest time and resources in an optimum course for professional development — one that is relevant to your job and aligned with your organization's business objectives.

Instead, it's tempting to follow the latest buzz or ignore the relevance of what they learn. Will you become more effective thanks to what you've learned, or does your development end once the training is over?

Such an ad-hoc approach to continuous professional development creates various issues, such as problematic progress tracking and gap measurement. Learning Journeys address these problems.

Benefits of Learning Journeys

The journeys are designed to overcome the drawbacks of working independently with little guidance:

(1) Applicability is high. Learning Journeys are designed by our expert consultants with working experience in the field, to ensure relevance to the participant's role and business needs.

(2) Learned material is applied and tested in practice through use cases. For example, an aspiring data scientist will be working on relevant use cases.

(3) Measuring impact on the workforce is much easier since employees follow similar paths. Small variations average out, and a clear picture emerges.

(4) Designed by the business to satisfy current and future needs, Learning Journeys are a valuable investment in individual careers and in the company as a whole.

(5) Current and potential employees can advance their careers through carefully crafted learning paths.

We have designed Learning Journeys for analytics translators, data scientists, and data engineers. You can find each Learning Journey in this guide:

» Analytics Engineering Learning Journey – Page 7
» Analytics Translation Learning Journey – Page 11
» Data Analyst Learning Journey – Page 16
» Data Scientist Learning Journey – Page 21
» Machine Learning Engineering Journey – Page 43
» Data Engineer Learning Journey – Page 59

Go to the academy ›

Advice and Inquiries

Need advice on which course is right for you? Curious about setting up the perfect data and AI in-company program?
Our director of sales, Marcel van Denderen, is available to answer your questions. Get in touch at +31 6 3623 7853 (phone or WhatsApp), or send an email: marcel.vandenderen@xebia.com

Machine Learning Engineering Learning Journey

Machine Learning Engineering

Following the boom of data and AI, one of the biggest challenges companies face is moving from early-stage concepts to robust applications that can actually deliver value in production. Solving this challenge requires having people that combine strong engineering skills with a keen affinity for machine learning: Machine Learning Engineers. However, with the explosion in available machine learning technologies and frameworks, becoming a Machine Learning Engineer is far from trivial. That's why we've designed several Machine Learning Engineer learning journeys that focus on key concepts and core technologies, providing ideal starting points for aspiring Machine Learning Engineers or more advanced Machine Learning Engineers that wish to deep dive into specific subjects.

Learning Journey

How do you become a machine learning engineer? Start here! We've put together a carefully crafted learning journey for data engineers. Knowing engineers love to figure things out on their own, we packed the program with opportunities to learn, hands-on, by solving real-life situations. Plus, there's plenty of practical philosophy, too.

Focusing on the fundamentals, we'll teach you how to build well-structured, production-ready machine learning applications in Python and how to run these using Docker and basic cloud services. Moving towards more advanced topics, we'll dive deeper into how to use MLOps practices to automatically deploy and retrain ML applications in the Cloud and on Kubernetes, using CI/CD and infrastructure-as-code to make sure our full stack is reproducible. Finally, in our specializations we do deep dives into specific technologies such as Airflow, dbt, Spark, etc. to make you the go-to expert in these powerful tools.

Learning Goals of the Fundamentals for Machine Learning Engineering

- Understand the fundamentals of machine learning engineering
- Build well-structured machine learning products in Python with guidance
- Know how containerization works and what it simplifies

Learning Goals for Advanced Machine Learning Engineers

- Design, build and deploy robust end-to-end machine learning solutions
- Implement machine learning patterns and MLOps strategies on Kubernetes or the Cloud
- Manage and deploy infrastructure-as-code using Terraform

Learning Goals for Machine Learning Engineers who want to specialize

- Become the go-to expert in specific topics such as deep learning, streaming data and experimentation
- Dive deep into specialized technologies such as Airflow, DBT and Spark and understand how these tools can super-charge your work

Machine Learning Engineering Learning Journey

Data Science with Python foundation → Production Ready ML → Docker → Cloud Fundamentals: Google Cloud, AWS or Microsoft Azure

Machine Learning Engineer Learning Journey

Machine Learning in the Cloud: Google Cloud, AWS or Microsoft Azure → Kubernetes → ML Ops → Data Science with Spark

Machine Learning Engineer SpecializationsLearning Journey

A/B Testing and Experimetns → Machine Learning Explainability → Spark Optimization and Tuning → dbt Learn → Streaming architecture at scale → Deep Learning → Apache Airflow

Machine Learning Engineering Journey

Certified Data Science with Python

3 days Fundamentals

Dive into the world of data science through data analysis and machine learning.

Do you know how to perform basic data analysis in Python, but want to take your skills to the next level? During this three-day data science deep dive, you will learn how to unlock Python's potential for data analysis and machine learning. This course covers training models with scikit-learn and best practices for transforming your data with pandas, with a perfect combination of both theory and practice.

This training IS for you if...

- You have some basic experience with Python (or any other programming language).
- You want to receive a certification in the field of data science.
- You are eager to get more out of your data and build your own predictive models.
- You would like to apply best practices to your data science projects.
- You hope to better communicate and collaborate with your data science colleagues.

This training is NOT for you if...

- You've never worked with Python or any other programming language before (check out our Python for Data Analysts training instead).
- You are already an experienced data scientist and you want to develop your skills further (check out our Advanced Data Science training and other specialized topics instead).
- You think data science is all hype and can't add value (check out our Analytics Translation training instead).
- You are not interested in practical applications, only the math behind the algorithms.

After the training, you'll be able to

- Perform exploratory data analysis on your datasets with pandas
- Train and evaluate machine learning models with scikit-learn
- Identify the right machine learning algorithm and metric for your data problem
- Prepare complex data for machine learning with techniques such as scaling, encoding, and imputing
- Apply best practices for data wrangling and model building

Course Topic Details

Data Wrangling with Pandas

- Fetch descriptive summary statistics of your data with simple operations
- Effectively select and filter parts of your data with loc
- Retrieve advanced statistics with groupby aggregations
- Extend your dataset by creating new columns with assign
- Structure your code neatly by chaining methods
- Learn to apply best practices, such as pipe and lambda, to prevent bugs

Machine Learning with Scikit-Learn

- Use scikit-learn to train classification and regression models
- Evaluate trained models with train/test set split and scikit-learn metrics

- ✓ Use scikit-learn transformers for categorical variable encoding, scaling and missing values imputation
- ✓ Pre-process complex data in scikit-learn with ColumnTransformer and Pipeline
- ✓ Tune pre-processing and model hyperparameters with gridsearch

Machine Learning Theory

- ✓ Identify the type of machine learning task (classification or regression, supervised or unsupervised, and others)
- ✓ Differentiate between several machine learning algorithms (such as linear regression, decision tree, support vector machine)
- ✓ Create models that generalize (underfitting and overfitting, train-test split, k-fold cross-validation)
- ✓ Understand how to evaluate your model's effectiveness with various metrics (such as precision and recall, F1, root mean squared error, r2)

Program

Day 1

- ✓ Master exploratory data analysis with Pandas
- ✓ Introduction to machine learning (theory)
- ✓ Build your first machine learning model on a real dataset

Day 2

- ✓ Learn about different machine learning algorithms and how to choose the best one
- ✓ Master complex data with advanced data pre-processing techniques to build strong predictive models
- ✓ Apply best practices for model building with scikit-learn pipelines

Day 3

- ✓ Supercharge your machine learning models with hyperparameter tuning
- ✓ Review material + exam

Or

- ✓ Put your gained skills into practice with a hackathon!

More info and dates ›

Machine Learning Engineering Journey

Production-Ready Machine Learning

2 days Advanced

Learn the best practices to bring your machine learning models into production.

This practical 2-day training prepares you to bring your machine learning models into production by using best practices. For example, you will learn how software engineers manage their code and the advanced Python features that can make your life easier. You will also learn how to go from notebooks to packages, and much more.

This training IS for you if...

- ✓ You want to move your code from an experimental Jupyter notebook to a mature Python package.
- ✓ You want to enhance the quality of your code and apply the current industry-standard tooling.
- ✓ You want to be able to collaborate better on projects with your colleagues.

This training is NOT for you if...

- ✗ You are comfortable just experimenting in Jupyter notebooks and not interested in evolving your projects so they can bring value to the organization.
- ✗ You are looking to enhance your machine learning knowledge (see CDSwP/Advanced DS).
- ✗ You want to go in-depth about data pipelines or deploying on specific cloud environments.

After the training, you'll be able to

- ✓ Create a high-quality Python package for your machine learning project that is easy to share, collaborate on, and deploy
- ✓ Write robust Python code that is easy to extend, debug, monitor and test
- ✓ Understand the importance of and what it means for a project to be *production-ready*
- ✓ Serve your models with an API or command-line interface

Course Topic Details

Best Practices

- ✓ Python version management, package managers, and virtual environments
- ✓ Characteristics of high-quality and maintainable code
- ✓ Automatic linting and code formatting with black, flake8, isort
- ✓ Ensure quality checks on every commit with pre-commit
- ✓ Type hinting and type checking with mypy

Quality Python package

- ✓ What to test in a data science project with pytest
- ✓ Effective logging and monitoring with logging
- ✓ Build beautiful documentation with Sphinx and MyST
- ✓ Create a command-line interface to your package with typer
- ✓ Build an API with the modern, fast high-performance web-framework FastAPI

Program

Day 1

- ✓ What makes a 'production-ready' project?
- ✓ From Jupyter notebook to Python package
- ✓ How to enhance the quality and robustness of your code

Day 2

- ✓ Quality Python package with documentation, logging and testing
- ✓ Static model serving patterns with CLI and API
- ✓ Capstone project: from notebook to package

More info and dates ›

Machine Learning Engineering Journey

Docker

1 day Fundamentals

Learn how to containerize your solutions.

Containerized solutions are here to stay — but what exactly are containers? And how do you build, use and run them? This course covers those questions and more. In a single day, you will learn exactly how Docker containers work — how to build them, how to place your solutions in them, and most importantly, how to run them.

After the training, you'll be able to

- ✓ Containerize your solution
- ✓ Interact with the container
- ✓ Properly structure the containerized solution

Course Topic Details

Best Practices

- ✓ How to work with volumes and containers
- ✓ How Docker builds layers and how this knowledge improves your Dockerfile
- ✓ Interacting with the containers
- ✓ The Joy of automation with Docker-compose
- ✓ How to automatically lint Dockerfiles

Program

- ✓ How to work with volumes and containers
- ✓ How Docker builds layers and how this knowledge
- ✓ Interacting with the containers
- ✓ How to automatically lint Dockerfiles
- ✓ How to enjoy the automation that docker-compose gives you

More info and dates ›

Machine Learning Engineering Journey

GCP Fundamentals: Core Infrastructure

1 day Fundamentals

Get to know the fundamentals of Google Cloud for big data and machine learning.

Before working with Google Cloud Platform (GCP), it is important to have the right skills and knowledge. This 1-day Foundation level training is all you need to get started!

The training offers a combination of presentations, demos, and hands-on labs. You will learn the value of Google Cloud Platform and how to incorporate cloud-based solutions into business strategies. We explore essential concepts and terminology for working with Google Cloud Platform (GCP). You will also get to know many of the computing and storage services available in Google Cloud Platform, including Google App Engine, Google Compute Engine, and Google Container Engine

Course Topic Details

Introduction to Google Cloud Platform

- ✓ Explain the advantages of Google Cloud Platform
- ✓ Define the components of Google's network infrastructure, including: Points of presence, data centers, regions, and zones
- ✓ Understand the difference between Infrastructure-as-a-Service (IaaS) and Platform-as-a-Service (PaaS)

Getting Started with Google Cloud Platform

- ✓ Identify the purpose of projects on Google Cloud Platform
- ✓ Understand the purpose of and use cases for Identity and Access Management
- ✓ List the methods of interacting with Google Cloud Platform
- ✓ Lab: Getting Started with Google Cloud Platform
- ✓ Google Cloud Identity and Access Management (IAM)
- ✓ The several technologies used to connect with GCP
- ✓ Cloud Launcher
- ✓ Cloud Deployment Manager

Google Compute Engine and Networking

- ✓ Identify the purpose of and use cases for Google Compute Engine
- ✓ Understand the basics of networking in Google Cloud Platform
- ✓ Lab: Deploying Applications Using Google Compute Engine

Google Cloud Platform Storage Options

- ✓ Understand the purpose of and use cases for: Google Cloud Storage, Google Cloud SQL, and Google Cloud Bigtable
- ✓ Learn how to choose between the various storage options on Google Cloud Platform
- ✓ Lab: Integrating Applications with Google Cloud Storage

Google Container Engine

- ✓ Define the concept of a container and identify uses for containers
- ✓ Identify the purpose of and use cases for Google Container Engine and Kubernetes
- ✓ Introduction to Hybrid and Multi-Cloud computing (Anthos)
- ✓ Lab: Deploying Applications Using Google Container Engine

Google App Engine and Google Cloud Datastore

- ✓ Understand the purpose of and use cases for Google App Engine and Google Cloud Datastore
- ✓ Contrast the App Engine Standard environment with the App Engine Flexible environment
- ✓ Understand the purpose of and use cases for Google Cloud Endpoints
- ✓ Lab: Deploying Applications Using App Engine and Cloud Datastore

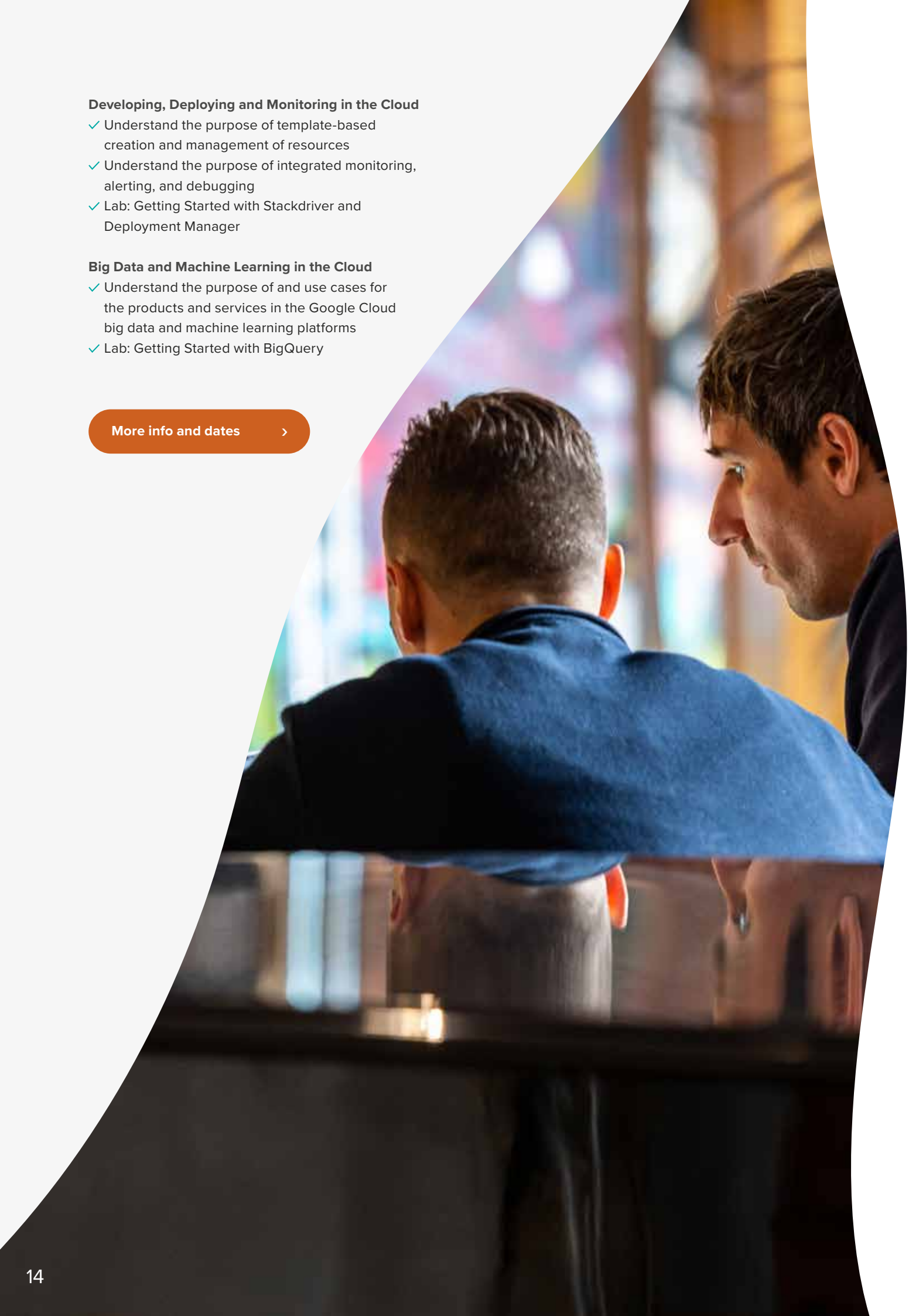
Developing, Deploying and Monitoring in the Cloud

- ✓ Understand the purpose of template-based creation and management of resources
- ✓ Understand the purpose of integrated monitoring, alerting, and debugging
- ✓ Lab: Getting Started with Stackdriver and Deployment Manager

Big Data and Machine Learning in the Cloud

- ✓ Understand the purpose of and use cases for the products and services in the Google Cloud big data and machine learning platforms
- ✓ Lab: Getting Started with BigQuery

More info and dates ›



Machine Learning Engineering Journey

AWS Cloud Fundamentals

1 day Fundamentals

Get to know the fundamentals of AWS.

You are going to be working in the Cloud soon, and AWS is the chosen platform. Now it's |time to gain the knowledge and skills to get started! This Foundation level training course provides you with an overview of AWS's most essential products and services. Following this training is also a great way to prepare for the AWS Cloud Practitioner exam.

AWS Foundation is perfect for

AWS Foundation is perfect for a technical crowd with none or limited AWS experience. You do not need to be familiar with Public Cloud to take part in the training. We start from scratch and share the basics of the platform with you. If you would like to explore AWS at an entry level before you participate, take a look at the AWS Cloud Practitioner Essentials: a digital course offered free of charge by AWS.

After the training, you'll be able to

- ✓ Define what the AWS Cloud is and the basic global infrastructure
- ✓ Describe basic AWS Cloud architectural principles
- ✓ Describe the AWS Cloud value proposition
- ✓ Describe key services on the AWS platform and their common use cases (for example, compute and analytics
- ✓ Describe basic security and compliance aspects of the AWS platform and the shared security model
- ✓ Define the billing, account management, and pricing models
- ✓ Identify sources of documentation or technical assistance (for example, whitepapers or support tickets)
- ✓ Describe basic/core characteristics of deploying and operating in the AWS Cloud

Course Topic Details

Introduction to Cloud Computing with AWS

- ✓ Services: VPC, EC2
- ✓ The meaning of Serverless
- ✓ Data and Storage options
- ✓ Networking
- ✓ IAM
- ✓ S3
- ✓ CloudWatch

Working with AWS

- ✓ Launching a server
- ✓ S3
- ✓ IAM
- ✓ SDK
- ✓ CLI
- ✓ CloudFormation

More info and dates ›

Machine Learning Engineering Journey

Certified Microsoft Azure Fundamentals (AZ-900)

1 day Fundamentals

Start working with Microsoft Azure!

To start working with Microsoft Azure right away, you have to know the ins and outs of the technology first. This 2-day fundamentals training will get you up and running in no time! You'll learn all about Azure itself, about Total Cost of Ownership and SLAs, but also about the many Azure cloud products and services. Includes the one-time taking the Microsoft AZ-900 exam!

To start working with Microsoft Azure right away, you have to know the ins and outs of the technology first. This 2-day fundamentals training will get you up and running in no time! You'll learn all about Azure itself, about Total Cost of Ownership and SLAs, but also about the many Azure cloud products and services. Includes the one-time taking the Microsoft AZ-900 exam!

This Azure Fundamentals training is perfect for

This training provides an "Azure Quickstart", an accelerating foundation for anyone who wants to get started with Microsoft Azure. Perfect for DevOps & AI Engineers, Developers, Database Administrators, Software and Solution Architects and Data Scientists who want to learn all about Azure's features and get hands-on practice. Experience with Microsoft's Cloud Computing Platform is not required, but technical knowledge regarding PowerShell scripting, a (random) programming language or tools like Visual Studio Code will help you to get the most out of this training.

After the training, you'll be able to

- ✓ Define what the AWS Cloud is and the basic global infrastructure
- ✓ Describe basic AWS Cloud architectural principles
- ✓ Describe the AWS Cloud value proposition
- ✓ Describe key services on the AWS platform and their common use cases (for example, compute and analytics)
- ✓ Describe basic security and compliance aspects of the AWS platform and the shared security model
- ✓ Define the billing, account management, and pricing models
- ✓ Identify sources of documentation or technical assistance (for example, whitepapers or support tickets)
- ✓ Describe basic/core characteristics of deploying and operating in the AWS Cloud

Course Topic Details

Azure Core Concepts

- ✓ A comprehensive overview of the fundamental concepts
- ✓ The architecture components
- ✓ Features and possibilities

Azure Core Services

- ✓ Azure Database
- ✓ Azure Compute
- ✓ Azure Storage
- ✓ Azure Networking

Module 3: Azure Core Solutions and Management Tools

- ✓ AI machine learning
- ✓ IoT
- ✓ Azure DevOps
- ✓ Monitoring and Management
- ✓ Serverless computing fundamentals

Module 4: Azure Security

- ✓ How do you protect your organization from threats in the Azure environment?
- ✓ How do you ensure secure network connectivity?

Module 5: Identity, Governance, Privacy and Compliance

- ✓ Identity management
- ✓ Various privacy services
- ✓ Azure's compliance
- ✓ Data protection standards

Module 6: Azure Cost Management and SLA's

- ✓ Manage Azure costs
- ✓ SLAs and service lifecycles

More info and dates ›

Machine Learning Engineering Journey

Google Cloud Platform Fundamentals: Big Data and Machine Learning

1 day Advanced

Before working with Google Cloud Platform (GCP), it is important to have the right skills and knowledge. This 1-day Foundation level training is all you need to get started with Big Data and Machine Learning on GCP.

The training offers a combination of presentations, demos, and hands-on labs that introduce you to the Big Data and Machine Learning capabilities of Google Cloud Platform (GCP). You will learn to process Big Data at scale for analytics and Machine Learning. You will explore the fundamentals of building new machine learning models and creating streaming data pipelines and dashboards.

After the training, you'll be able to

- ✓ Get started with Big Data and Machine Learning in the Google Cloud
- ✓ Have a clear overview of the Google Cloud Platform
- ✓ Understand the data processing capabilities

Program

Introducing Google Cloud Platform

- ✓ Google Platform Fundamentals Overview
- ✓ Google Cloud Platform Big Data Products

Compute and Storage Fundamentals

- ✓ CPUs on demand (Compute Engine)
- ✓ A global filesystem (Cloud Storage)
- ✓ CloudShell
- ✓ Lab: Set up a Ingest-Transform-Publish data processing pipeline

Data Analytics on the Cloud

- ✓ Stepping-stones to the cloud
- ✓ Cloud SQL: your SQL database on the cloud
- ✓ Lab: Importing data into CloudSQL and running queries
- ✓ Spark on Dataproc
- ✓ Lab: Machine Learning Recommendations with Spark on Dataproc

Scaling Data Analysis

- ✓ Fast random access
- ✓ Datalab
- ✓ BigQuery
- ✓ Lab: Build machine learning dataset

Machine Learning

- ✓ Machine Learning with TensorFlow
- ✓ Lab: Carry out ML with TensorFlow
- ✓ Pre-built models for common needs
- ✓ Lab: Employ ML APIs

Data Processing Architectures

- ✓ Message-oriented architectures with Pub/Sub
- ✓ Creating pipelines with Dataflow
- ✓ Reference architecture for real-time and batch data processing

Summary

- ✓ Why GCP?
- ✓ Where to go from here
- ✓ Additional Resources

More info and dates ›

Machine Learning Engineering Journey

Big Data on AWS

3 days Advanced

In this training, you will learn about cloud-based Big Data solutions such as Amazon EMR, Amazon Redshift, Amazon Kinesis, and the rest of the AWS Big Data platform. We will show you how to use Amazon EMR to process data using the broad ecosystem of Hadoop tools like Hive and Hue. We will also teach you how to create Big Data environments, work with Amazon DynamoDB, Amazon Redshift, Amazon QuickSight, Amazon Athena, and Amazon Kinesis, and leverage best practices to design Big Data environments for security and cost-effectiveness.

After the training, you'll be able to

- ✓ Define AWS data analytics services and understand how they integrate with each other
- ✓ Explain how AWS data analytics services fit in the data life cycle of collection, storage, processing, and visualization

This Big Data Training Is Perfect for

- ✓ Solutions architects
- ✓ SysOps administrators
- ✓ Data scientists
- ✓ Data analysts

Prerequisites

- ✓ At least 5 years of experience with data analytics technologies
- ✓ At least 2 years of hands-on experience working with AWS
- ✓ Experience and expertise working with AWS services to design, build, secure, and maintain analytics solutions

Program

Navigating the Big Data Landscape

- ✓ Fit AWS solutions inside a Big Data ecosystem
- ✓ Choose appropriate AWS data storage options
- ✓ Comprehend and manage costs and security for a Big Data solution
- ✓ Identify options for ingesting, transferring, and compressing data

Elastic Map Reduce

- ✓ Leverage Apache Hadoop in the context of Amazon EMR
- ✓ Identify the components of an Amazon EMR cluster, then launch and configure an Amazon EMR cluster
- ✓ Use common programming frameworks available for Amazon EMR, including Hive, Pig, and streaming
- ✓ Improve the ease of use of Amazon EMR by using Hadoop User Experience (Hue)
- ✓ Use in-memory analytics with Apache Spark on Amazon EMR

Learn the Ins and Outs of Various Big Data Services

- ✓ Identify the benefits of using Amazon Kinesis for near real-time Big Data processing
- ✓ Leverage Amazon Redshift to efficiently store and analyze data
- ✓ Leverage Amazon Athena for ad-hoc query analytics
- ✓ Use AWS Glue to automate extract, transform, and load (ETL) workloads
- ✓ Use visualization software to depict data and queries using Amazon QuickSight

More info and dates ›

Machine Learning Engineering Journey

Microsoft Certified: Azure Data Scientist Associate

3 days Advanced

Candidates for the Azure Data Scientist Associate certification should have subject matter expertise in applying data science and machine learning to implement and run machine learning workloads on Azure.

Responsibilities for this role include designing and creating a suitable working environment for data science workloads; exploring data; training machine learning models; implementing pipelines; running jobs to prepare for production; and managing, deploying, and monitoring scalable machine learning solutions.

A candidate for this certification should have knowledge and experience in data science by using Azure Machine Learning and MLflow.

After the training, you'll be able to

- ✓ Design and prepare a machine learning solution
- ✓ Explore data and train models
- ✓ Prepare a model for deployment
- ✓ Deploy and retrain a model
- ✓ Automate machine learning model selection
- ✓ Explore differential privacy
- ✓ Explain machine learning models
- ✓ Detect and mitigate unfairness in models
- ✓ Monitor models & data drift
- ✓ Explore security concepts

Program

Create machine learning models

- ✓ Explore and analyze data with Python
- ✓ Train and evaluate regression, classification, clustering, and deep learning models

Microsoft Azure AI Fundamentals: Explore visual tools for machine learning

- ✓ Use Automated Machine Learning
- ✓ Create a regression, classification, and clustering model with Azure Machine Learning designer

Build and operate machine learning solutions with Azure Machine Learning

- ✓ Introduction to the Azure Machine Learning SDK
- ✓ Train a machine learning model
- ✓ Work with Data & Compute
- ✓ Orchestrate machine learning with pipelines
- ✓ Deploy real-time machine learning services and batch inference pipelines
- ✓ Tune hyperparameters

Build and operate machine learning solutions with Azure Databricks

- ✓ Get started with Azure Databricks
- ✓ Work with data in Azure Databricks
- ✓ Prepare data for machine learning
- ✓ Train a machine learning model
- ✓ Use MLflow to track experiments
- ✓ Manage machine learning models
- ✓ Track Azure Databricks experiments
- ✓ Deploy Azure Databricks models
- ✓ Tune hyperparameters
- ✓ Distributed deep learning with Horovod and Azure Databricks

More info and dates ›

Machine Learning Engineering Journey

Kubernetes

1 day Advanced

Learn how to deploy your containerized solutions in a container-based cluster.

Containerized solutions are here to stay — but how to deploy them? In a single day, you will learn how to deploy your containerized solutions in a Kubernetes cluster and manage all aspects of its life cycle.

After the training, you'll be able to

- ✓ Deploy your container(s) to Kubernetes
- ✓ Manage the solution and the cluster

Course Topic Details

- ✓ Kubernetes core concepts
- ✓ Configuration
- ✓ Multi-container pods
- ✓ Observability
- ✓ Pod design
- ✓ Services and Networking
- ✓ State persistence
- ✓ Introduction to Helm

Program

- ✓ Kubernetes core concepts
- ✓ Configuration and Multi-container pods
- ✓ Observability and Pod design
- ✓ Services and Networking
- ✓ State persistence
- ✓ Introduction to Helm

More info and dates ›

Machine Learning Engineering Journey

MLOps Training

2 days Advanced

During this training, you will learn how to apply MLOps principles such as continuous training, continuous deployment and end-to-end monitoring to build end-to-end solutions in one of the public clouds (AWS, Azure or GCP).

This training IS for you if...

- ✓ Already have a solid understanding of ML, and want to take your models outside of the development phase.
- ✓ Already have basic SWE skills (Basic understanding of Docker, Python, Git).
- ✓ Want to incorporate best practices from Software Engineering.
- ✓ Want to learn more about the Cloud.

This training is NOT for you if...

- ✗ Want to learn more about developing ML models itself (this knowledge is already assumed).
- ✗ Do not have basic programming experience. In that case, an introductory course is advised.
- ✗ Are mainly interested in doing (exploratory) research. This course is much oriented towards ML engineering.

After the training, you'll be able to

- ✓ Have a solid understanding of all the necessary components in an ML system. Including best practices, common design challenges, etc.
- ✓ Create a Machine Learning Pipeline In AzureML.
- ✓ Deploy your model on Azure as scalable API on Azure Container Instances
- ✓ Integrate and deploy all code through a CI/CD pipeline in Azure DevOps

Course Topic Details

Best Practices

- ✓ Key principles of MLOps
- ✓ How to design solutions for deploying models to the cloud
- ✓ How to build an ML pipeline for reproducibly training a model
- ✓ How to automatically deploy and schedule ML pipelines for automated re-training
- ✓ How to track models and metrics in a model library
- ✓ How to deploy and monitor an ML model as scalable REST APIs
- ✓ Practical advice on how productionize ML models in the cloud

Program

Day 1

- ✓ Key MLOps principles
- ✓ Creating a solution design
- ✓ Building an ML pipeline
- ✓ Deploying an ML pipeline with CI/CD

Day 2

- ✓ Scheduling an ML pipeline for automated training
- ✓ Tracking trained models and their metrics
- ✓ Deploying models as REST APIs

More info and dates ›

Machine Learning Engineering Learning Journey

Specializations in Machine Learning Engineering

Want to become the go-to expert in specific topics such as deep learning, streaming data and experimentation? Want to deep dive into specialized technologies such as Airflow, DBT, and Spark and understand how these tools can super-charge your work? Then take one of our following training courses.

A/B testing and experimentation

In business, how can you know if changing something will improve it or not? The two-day A/B Testing and Experiments training will teach you everything you need to successfully run your own experiments.

More info and dates ›

DBT

Own the workflow! Learn everything you need to know to improve your dbt (data build tool) skillset. Learn about modeling strategies, automated testing, and performance optimization.

More info and dates ›

Apache Airflow

Learn to author, schedule, and monitor workflows through hands-on experience with the leading open source platform in the space.

More info and dates ›

Deep Learning

This three-day course covers the fundamentals of Deep Learning and explains how it differs from traditional Machine Learning. You will gain the practical skills you need to implement your own Deep Learning algorithms and learn to apply them to unstructured data.

More info and dates ›

Apache Spark – Data processing at scale

This training goes deep down into one of the most popular and scalable tools in the market for large-data transformation: Apache Spark!

More info and dates ›

Machine learning explainability

Get a toolbox of interpretability techniques that you can use in your daily work, understanding when and how you can and should use it.

More info and dates ›

Apache Spark – Optimization and tuning best practices

Processing data efficiently can be challenging as it scales up. Building up from the experience we built at the largest Apache Spark users in the world, we give you an in-depth overview of the do's and don'ts of one of the most popular analytics engines out there.

More info and dates ›

Streaming architecture at scale

While some apps enjoy the luxury of processing in a batch oriented fashion, others, as in the IoT ecosystem, expect events to be ingested and processed as they occur. This training focuses on two key players on the streaming-side of data processing: Apache Kafka and Apache Spark!

More info and dates ›

In-company Data and AI Learning Journeys

Do you want to tackle data issues with the people who understand them best? With our in-company programs, you learn from expert trainers and your colleagues. We give your organization the most effective training programs so your employees can be even more successful data professionals.

Our Approach

A successful in-company program depends on more than the course itself. With our four-step approach, we ensure that your organization achieves maximum value.

(1) Intake

We conduct an intake to discuss the project, challenge, or issues you want to tackle. During the intake, we also explore any ideas you may already have about the execution of the in-company program.

(2) Advice and offer

Based on your needs, our team gets to work in creating a comprehensive program. Together, we refine and adjust it.

(3) Execution

We discuss the specifics of the program with the trainer and also send an intake form to all participants so we know their knowledge level and learning goals. Our academy coordinator provides the necessary course material and participants improve their skills and gain knowledge through theory, hands-on labs, and discussions.

(4) Follow-up

After the training, we conduct an extensive evaluation and use this input to improve the program. We continuously update our courses to ensure that they are current with the latest information, techniques, and technologies.

After the program, we can support your organization and your employees in getting even more out of the courses. This can be achieved through coaching, follow-up days, or more in-depth and specialized additional courses.

The 7 Benefits of In-Company Courses

(1) Adapted to your situation

With an in-company program, we can adjust the course material to your needs, challenges, and specific situation. This makes the course even more applicable.

(2) Extensive offering of over 25 courses

All our public classes are available as in-company courses. Next to that, we have a host of courses exclusively available for in-company programs. Do you have a data and AI challenge but don't see the right course? With our varied expertise, we can shape the right course for you.

(3) Best trainers

All our trainers are practitioners in the field. Through their work as consultants, they solve challenges like yours every day.

(4) You decide on the right place and time

You decide what works best for your employees — in-house or external. Want a specific date? No problem, it's up to you.

(5) Financially attractive

Our in-company programs are available starting at six participants, making them cost-efficient.

(6) High quality, with an average score of 8.7

Participants in our courses are highly satisfied with the quality of their learning.

(7) Intensive guidance throughout the process, from start to finish

We collaborate closely to shape the program you need. Our skills advisors, academy coordinators, and trainers support you every step of the way.

From Team to Boardroom – Data and AI Courses for Your Entire Organization

Embedding data-driven decision-making in your organization by leveraging AI products needs more than just a good data team. You cannot become fully data-driven without executives sponsoring use cases and leading by example. Projects fall short when the business challenge is not connected to the technical one by data science product owners. Your general workforce requires to be open to data and AI and knowledgeable about the changes they bring. That's why we provide programs for your entire workforce.

C-Level and Executives

There can be no project without buy-in, and to have buy-in, your C-level and other executives need to understand the possibilities and pitfalls of AI. Our programs provide them with a deep understanding of data and AI, but also what it takes to generate successful use cases and their role in sponsoring these. This way, AI solutions can be developed more quickly, but also with a better focus on value generation.

Data Engineers and Data Scientists

We deepen and widen the skills of your data engineers and data scientists. With our programs focused on Python, Spark, Airflow, and Docker, they will increase their ability to effectively and efficiently use these critical techniques.

With courses such as Deep Learning, A/B Testing, Time-Series, and Recommenders, they acquire more specialized knowledge and skills.

Data Science Product Owners

Whether you call them data science product owners, analytics translators, business translators, or AI consultants—they all support the connection between data teams and the business.

We teach these product owners how to generate AI use cases, check the value of a project every step of the way, and ensure the business fully adopts and uses AI solutions.

General Workforce

Your general workforce might be the most critical part of your data-driven transformation. Without support from the business, there can be no success in data and AI. They are the people that will drive the adoption of the data products and ultimately change the way you do business.

Empowering them with knowledge of data and AI and informing them about the possibilities creates analytics ambassadors throughout your entire organization. We do this with online courses, workshops, and board games, to name a few of the options. Employees will be able to generate powerful ideas on what data and AI products can enhance their work and how to work with data science product owners on generating use cases.

In-company Courses and Programs

Are you looking for a course to teach your data engineers more about Apache Airflow? Or do you need a full program to take your organization's data scientists from juniors to seniors? At Xebia, we have the right expertise to offer the flexibility you need. We are happy to show you some examples of our work.

Data Engineering and Data Science Accelerator – ING

For ING, we executed the full data engineering and data science learning journeys described earlier in this guide. With a 13-month/13-modules program, ING's data professional went through an entire journey from junior to senior. Each of the 13 modules featured one day of theory and one day of hands-on labs. The program's breadth covered the full spectrum of relevant knowledge for data engineers and data scientists.

Data Science Specialization Training – ABN AMRO

ABN AMRO's wide variety of products and departments creates a training challenge. To ensure applicability and benefits for all of its data scientists, the training program was set up as a set of single-day courses focusing on several data science specializations, including reinforcement learning, recommender systems, and data science engineering.

Data Discovery Workshops – Schiphol and ING

To educate C-level executives, managers, and the general workforce on the "art of the possible" for data and AI, Xebia regularly hosts Data Discovery Workshops. In these workshops, participants get to work on creating data and AI use cases they can apply in their work. For ING, we organized these workshops as a worldwide program for their C-level executives. At Schiphol, we joined the internal innovation week, showing managers and employees what they can achieve with data and AI.

International Data Science Program – DSM

Chemical enterprise DSM asked Xebia to design an international program to enhance its internal data science capabilities. With courses taught in the Netherlands, the United States, and Switzerland, the program contained foundational, intermediate, and advanced courses on Python. Some of the classes came from Xebia's standard offerings, while others were created especially for DSM.

Apache Airflow Ask Me Anything – Bosch

Bosch Germany was searching for possibilities to improve Apache Airflow expertise within the company. As Bosch's data engineers already possessed solid knowledge of Apache Airflow, they did not just want a training course. Bosch asked Xebia for the best approach, and together we decided to send our most senior Apache Airflow expert. This expert was on-site at Bosch for two days to answer all questions about Apache Airflow and give hands-on support on active projects.

Advanced Apache Spark Users – Schiphol

In addition to training programs, Xebia worked with Schiphol on a data engineering consultancy. During this consultancy, Schiphol employees learned the finer details of Apache Spark and were eager to further improve their skills. Xebia, therefore, provided Schiphol's data engineers with an Advanced Apache Spark course, going far beyond the standard Data Science with Spark course.

Supporting bol.com as a Frontrunner in Data and AI

Customer Challenge To enhance in-house data engineering and data science capabilities and remain a frontrunner in data and AI.

Our Solution An extensive in-company training program comprised of 15 to 20 courses per year, running since 2018.

Outcome Professionals working at bol.com maintain up-to-date skills with the latest data engineering and data science technologies. They can apply the knowledge they acquire in training to their day-to-day work, supporting bol.com's business.

Data for Decision-Making First

Founded in 1999 as a pure internet company, bol.com doesn't have the baggage of traditional organizations in becoming data-driven. The company's advanced analyst Melissa Perotti explained, "Decision-making at bol.com is based on data first. The data doesn't confirm someone's hunch; it tells us where to go."

As the largest e-commerce platform in The Netherlands and Belgium, bol.com has vast amounts of data for decision-making at its fingertips. "We use data science in many places—our search engine and recommender systems, forecasting of sales and customer service, and our chatbot. We have a mature data landscape and are proud of what we have achieved so far—and we're ready to take it all the next level."

Over the past two years, bol.com's data science capability quickly matured and now the shopping platform's next challenge is to extend this beyond its data engineers and data scientists.

Great Fit with Xebia

For the data engineering and data science training program, bol.com looked for a partner to support their journey. The two main criteria for its program were applicability and being up-to-date with the newest theories and models. After comparing different partners, bol.com chose Xebia. "Xebia understood how we saw our data science capability, and where we wanted to go," said Perotti. "It was also really flexible in setting up the programs, so the courses were perfectly tailored to our situation, and the material was very applicable in our work."

A Wide Host of Courses

The in-company training program offered a wide range of courses to data professionals at bol.com. Courses included programs for data science starters like Python for Data Analysts, but also more advanced courses such as Deep Learning.

Most of the courses are part of the curriculum Xebia offers and were adapted to fit bol.com's needs. For example, the A/B Testing and Experiments course was split into a technical and a non-technical training. The technical training focused on data scientists and ways they could apply experimenting in their day-to-day work. The non-technical training was aimed at managers, to give them an insight in the possibilities.

Extending the Training Program

After providing several courses for bol.com in 2018 and 2019, the program has been extended into 2020 with several additional courses.

Xebia has been heavily involved in setting up and supporting the program. "What is great about the program with bol.com is that we really work together on the execution," he explained. "After every training, we conduct an extensive evaluation to see if there are things we can improve from either side."

Data and AI Accelerator Program for Enterprises

Transforming into a Data-Driven Global Enterprise

As an enterprise, how do you become fully data- and AI-driven? Ingraining data and AI into the DNA of your employees' way of working is an immense challenge. We can help you climb that mountain.

Gain Hands-On Experience with Data and AI

In collaboration with you, GoDataDriven shapes and executes this program, from baseline assessment to delivering training and executing AI use cases.

What makes the program so effective and unique is the combination of learning methods. It's more than a classroom or online course. All participants go hands-on with data and AI and create their own use cases.

For the Entire Workforce

The Data and AI Accelerator Program is specially designed for enterprise organizations that want to upskill their entire workforce. It offers learning journeys, courses, and certifications for every type of employee.

Whether it's C-level executives, data scientists, or employees who rarely touch data, this program covers everyone's training needs.

Why Enterprises Love the Data and AI Accelerator Program

Program Design from Start to Finish

We design the full Data and AI Accelerator Program with you. The program's design is completely tailored toward your organization's goals, challenges, and capabilities.

Authority From the Field, Like No Other

All our trainers work in the field as consultants. They experience and solve challenges like yours every day. Your organization will benefit enormously from this experience. We don't come from an abstract background — we are practitioners as much as we are teachers.

High Satisfaction Rate

The Data and AI Accelerator Program at ING was an overwhelming success. Participants rated their training courses with an 8.5.

The program enabled ING to establish best practices and a standardized way of working, letting ING tackle AI challenges better than ever.

Learn Today, Apply Tomorrow

Applicability of our training courses is a core GoDataDriven value. Our programs are designed so your employees can apply their new knowledge immediately. In the Data and AI Accelerator Program, every trainee participates in data and AI use cases, giving them direct hands-on experience and a deeper understanding.

Five Elements of the Data and AI Accelerator Program

The Data and AI Accelerator Program is designed using a set framework. This allows for both quick implementation and flexibility to tailor the program to your needs.

(1) Assessment and Goal-Setting

We assess the baseline capabilities of your workforce, determine the personas, and set goals for the desired end result.

(2) Learning Journey Design

With the start and end points determined, a learning journey is designed for each persona. This learning journey includes all types of training, examination, and certification, together with governance and training logistics.

(3) Curriculum Creation

We review and update our existing material and create new material if required. To ensure the material fits with your employees, we field-test it with selected colleagues.

(4) Program Execution

Our trainers work on upskilling your employees to become truly data- and AI-driven. The program includes a mix of classroom and online training, meetups, and hackathons. We constantly update the course material, so your employees will always have the latest information.

(5) Evaluation and Improvements

After each session, participants fill out an evaluation. Based on their feedback, we evaluate the session and improve future sessions.

Securing Budget for Your Training Course

Throughout this guide, we have emphasized the necessity of staying up-to-date for data professionals. The Internet of Things, the growing power of cloud providers, ever-changing technologies, specialization—these are all valid and urgent reasons to develop your skills. However, these reasons are also rather generic. If you want to free up a training budget at your organization, what should be your approach? How do you make your case clear and convincing?

Need a helping hand in preparing your case? Get in touch with our director of sales, Marcel van Denderen, at marcel.vandenderen@xebia.com or +31 6 3623 7853.

Talk to HR and Your Colleagues

The best place to start is HR. They will be able to tell you whether there is a policy for the training budget and any requirements. Sometimes it's as simple as putting in a request or picking one of the options in your company's training portal.

Does your organization not offer a standard training budget, or is there no policy? In that case, ask your colleagues. If they have attended a course in the past, they can tell you what their approach was.

Do Your Research

If freeing up training budget requires you to convince your manager, it's vital to conduct a proper investigation into the different options. You may feel you will benefit from a three-day course, costing around €1,800, but your manager might have a very different view.

Research the various options, and be open-minded about what you find. A €20 Udemy course might not be a great option, but clarify for yourself why it's not. This will help you convincingly present your case.

Prepare Your Case

It's essential to present a well-prepared case that shows why your preferred training option is best. Start your case with the conclusion—explain the value your proposed course will bring to the company. How much money will it save, how much faster will you be able to deploy a model? Of course, your manager is interested in your development, but they also need to validate the expense. This is why it's important to emphasize its value for the company.

In addition to demonstrating the value, it's wise to show alternatives and how they relate to your preferred option. It may sound counterintuitive, but it's likely that your manager or HR will ask about potential alternatives. Being prepared and presenting sound reasoning for your choice only strengthens your case.

Show the Value

Validating your chosen course doesn't end when you receive approval. Before attending the course, have a clear plan for what you want to learn and how you intend to apply it. When returning to work, put your learnings into practice and show your manager the results you have generated and the value it has delivered for your organization.

Big Enough to Deliver, Small Enough to Care

At Xebia, we have been delivering data training courses since 2010. Over the past decade, we have dedicated a lot of time in laying the foundation for our unique curriculum and refining it with every iteration. Based on our own experiences, we are able to bring the practical skills professionals need to add value every day. We deliver training world-wide for global enterprises, but make sure to provide a personal experience for every participant. Not only with content and delivery methods that seamlessly fits their needs, but also by delivering training in many popular languages.

Our training numbers

460
companies served

1,412
days of training

18,080+
people trained

8.9/10
satisfaction rating

Our Track-Record

We are full stack
We organize, we build, we train. Everybody is talking about data and AI. Our team has been implementing scalable data solutions since 2009.

In-depth knowledge
We contribute to Apache Airflow, Apache Kafka, Keras, Scikit-learn, pandas, Apache Avro, Apache Parquet, Apache Flink, Apache Spark and many more.

Learn today, apply tomorrow
Battle-tested during more than 10,000 hours of classroom training, the extensive, hands-on curriculum has been developed with applicability in mind.

Each trainer is a consultant
We bring real knowledge from the trenches. Xebia has the expertise — built through years of working with the top enterprises in Europe.

Our Global Training Delivery Capabilities

Our international group of expert trainers delivers classes in the most popular European languages. Our trainers are native in Dutch, English, France, German, Italian, Spanish.

Trusted by Global Leaders



Trainer Overview


Alexander Bij


Cor Zuurmond


Daniel van der Ende


Herbert van Leeuwen


James Hayward


Juan Perafan Rodriguez


Julian de Ruiter


Kris Geusebroek


Lucy Sheppard


Lysanne Van Beek


Manolis Manousogiannis


Marysia Winkels


Nelli Gofman


Pádraic Slattery

Razvan Vacaru

Rens Dimmendaal

Roel Bertens

Rogier van der Geer

Steven van Duin

Tennyson Wu

Thomas van Latum

Timo Uelen

Travis Dent

Vadim Nelidov

Yke Rusticus

Xebia Academy Team for Data and AI Learning Journeys

Giovanni Lanzani – Managing Director Data
As managing director data of Xebia's Academy, Giovanni is responsible for the content of our courses and learning journeys. He also leads Xebia's own development efforts.

Marcel van Denderen – Sales Director
Marcel is responsible for all sales of the Xebia Data and AI Academy. If you want to improve the data and AI skills of your employees with one of our in-company programs, Marcel can advise you on the best setup.

Nico de Jonge – Head of Business Development
Nico is frequently in touch with organizations all over the world about their challenges in data literacy. If you have questions about upskilling your organization, team or self, Nico is happy to help you on the best way to move forward.

Diego Teunissen – Learning and Development Consultant
As a Learning and Development Consultant, Diego is responsible for the public courses on data and AI at Xebia Academy. The data field is dynamic like no other. Keeping your skills up to date is key to staying ahead. If you need advice on which training course fits your learning goals, Diego is always there to help you or your team further.

Giovanni Lanzani

Marcel van Denderen

Nico de Jonge

Diego Teunissen

Frequently Asked Questions

How can I get advice before booking a course?
Diego Teunissen, Learning and Development Consultant, or Marcel van Denderen, Director of Sales, will happily answer your questions and advise you on our courses. You can reach them via email, or by phone:
Diego Teunissen,
diego.teunissen@xebia.com,
+31 6 1591 4440
Marcel van Denderen,
marcel.vandenderen@xebia.com,
+31 6 3623 7853

I'm interested in attending a course that is not on the current schedule. What do I do?
We plan training courses throughout the entire year, so if you don't see the course of your choice on the schedule, there's still a chance we could add it. Get in touch with Marcel van Denderen to discuss t he possibilities via email, marcel.vandenderen@xebia.com, or by phone, +31 6 3623 7853.

Can I cancel a booked course?
If you are unable to attend a booked course, please send an email to academy@xebia.com. If you cancel up to one month before the first training day, there is no cancellation charge. Cancellations between one month and two weeks of the first day pay 50%. If you cancel a course within two weeks of the first training day, we charge the full price.

Do I receive a certificate after attending a course?
Yes, you will receive a Xebia Certificate of completion after completing a Xebia course.

Do I need to bring a laptop and install any software?
All of our courses include extensive hands-on sections requiring a laptop. The need to install software differs between courses. Please refer to the specific course information for details. You will also receive a confirmation email shortly before the training date that lists any requirements.

Training	Duration	Topic	Level	Page	More info at
A/B Testing and Experiments	2 days	Data Science	Fundamentals	27	gdd.li/ab-testing-exp
Advanced Analytics Translation	2 days	Analytics Translation	Advanced	14	gdd.li/advanced-analytics-translation
Advanced Data Science with Python	2 days	Data Science	Advanced	31	gdd.li/adv-ds
Advanced Power BI	2 days	Data Analysis	Specialty	20	gdd.li/da-advanced
Analytics for Executives	4 hours	Analytics Translation	Fundamentals	71	gdd.li/analytics-for-executives
AWS Cloud Fundamentals	1 day	Machine Learning Engineering	Fundamentals	45	gdd.li/cf-aws
Bayesian Statistics Training	2 days	Data Science	Fundamentals	28	gdd.li/bayesian
Big Data on AWS	3 days	Machine Learning Engineering	Advanced	14	gdd.li/aws-big-data
Certified Analytics Translation	2 days	Analytics Translation	Fundamentals	13	gdd.li/analytics-translation
Certified Apache Airflow	2 days	Data Engineering	Advanced	66	gdd.li/apache-airflow
Create Data Science Products	2 days	Data Engineering	Advanced	63	gdd.li/ds-products
Data Processing at Scale	3 days	Data Engineering	Fundamentals	48	gdd.li/data-scale
Data Science with Python (Certified)	3 days	Data Science and Machine Learning Engineering	Fundamentals	25/48	gdd.li/python-certified
Data Science with Spark	3 days	Data Science	Specialty	38	gdd.li/ds-spark
Data Visualization and Storytelling	2 days	Analytics Engineering and Data Science	Advanced	10/24	gdd.li/data-visualization
dbt Learn	3 half-days	Analytics Engineering and Data Engineering	Fundamentals	8/61	gdd.li/dbt-learn
Deep Learning	3 days	Data Science	Specialty	36	gdd.li/deep-learning
Deep Learning for Natural Language Processing	2 days	Data Science	Specialty	42	gdd.li/deep-learning-nlp
Docker	1 day	Machine Learning Engineering	Fundamentals	52	gdd.li/docker
Docker and Kubernetes	3 days	Data Engineering	Fundamentals	65	gdd.li/dockerkubernetes
Google Cloud Platform Fundamentals: Big Data and Machine Learning	1 day	Machine Learning Engineering	Advanced	53	gdd.li/gcp-big-data-ml
Google Cloud Fundamentals: Core Infrastructure	1 day	Machine Learning Engineering	Fundamentals	46	gdd.li/gcf-core-infrastructure
Innovation Through Experiments	1 day	Analytics Translation	Fundamentals	12	gdd.li/innovation-experiments
Kubernetes	1 day	Data Science and Machine Learning	Advanced	55	gdd.li/kubernetes

Machine Learning Explainability	1 day	Data Science	Advanced	34	gdd.li/ml-exp
Certified Microsoft Azure Fundamentals (AZ-900)	1 day	Machine Learning Engineering	Fundamentals	50	gdd.li/cf-azure
Microsoft Certified: Azure Data Scientist Associate	3 days	Machine Learning Engineering	Advanced	54	gdd.li/azure-data-science
Microsoft PL-300	3,5 days	Data Analysis	Advanced	18	gdd.li/power-bi-pl-300
MLOps Training	2 days	Machine Learning Engineering	Advanced	56	gdd.li/mlops-training
Optimizing Apache Spark and Tuning Best Practices	2 days	Data Science and Data Engineering	Specialty	40/68	gdd.li/optimizing-spark
Power BI in a Day	1 day	Analytics Engineering and Data Analyis	Fundamentals	9/17	gdd.li/da-power-bi
Production-Ready Machine Learning	2 days	Data Science and Machine Learning Engineering	Advanced	33/51	gdd.li/prod-ml
Python for Data Analysis	2 days	Data Science	Fundamentals	23	gdd.li/python-da
Python for Data Engineers	2 days	Data Engineering	Advanced	62	gdd.li/python-engineer
Streaming Architecture at Scale	2 days	Data Engineering	Specialty	70	gdd.li/arch-scale
Time Series Analysis and Forecasting	2 days	Data Science	Advanced	29	gdd.li/time-series


RHYME

Xebia | Data

Questions about our data and
AI learning journeys?
marcel.vandenderen@xebia.com
+31 6 3623 7853