Course : Certifying course Collecting, storing and making available data from an artificial intelligence project

Skills block of the RNCP 37827BC01 title

Practical course - 28d - 196h00 - Ref. ZRI

This training course represents the first block of skills in the state-recognized RNCP Level 6 (Bac +3) "Artificial Intelligence Developer" qualification. It will teach you how to automate data extraction, develop SQL queries, aggregate data, create RGPD-compliant databases, and develop APIs to share data. These skills are crucial in a world where effective data management is essential for decision-making and regulatory compliance.

Teaching objectives

At the end of the training, the participant will be able to:

- Automate data extraction
- Develop SQL queries to extract data from DBMS and big data systems
- Develop rules for aggregating data from different sources
- Creating a database in compliance with the RGPD
- Develop an API to make the dataset available

Intended audience

Anyone wishing to collect, store and make available data from an artificial intelligence project.

Prerequisites

Hold a level 5 diploma (Bac +2), with knowledge of object programming and SQL. If this is not the case, hold a level 4 diploma (BAC) and 3 years' experience in application development, subject to validation of the VAP file by the certifier.

PARTICIPANTS

Anyone wishing to collect, store and make available data from an artificial intelligence project.

PREREQUISITES

Hold a level 5 diploma (Bac +2), with knowledge of object programming and SQL. If this is not the case, hold a level 4 diploma (BAC) and 3 years' experience in application development, subject to validation of the VAP file by the certifier.

TRAINER QUALIFICATIONS

The experts leading the training are specialists in the covered subjects. They have been approved by our instructional teams for both their professional knowledge and their teaching ability, for each course they teach. They have at least five to ten years of experience in their field and hold (or have held) decision-making positions in companies.

ASSESSMENT TERMS

The trainer evaluates each participant’s academic progress throughout the training using multiple choice, scenarios, hands-on work and more.
Participants also complete a placement test before and after the course to measure the skills they’ve developed.

Certification

Le bloc de compétences est validé à travers une mise en situation. L'évaluation doit se faire dans un contexte de réalisation d'un service numérique réel ou fictif basé sur l'usage de données, à partir du cadrage pour la réalisation d'un service numérique (spécifications fonctionnelles et techniques par exemple). Le projet évalué a pour but d'optimiser, d'automatiser, de pérenniser et de mettre à disposition les flux de données et les données, utiles et nécessaires à la réalisation du service numérique, par les équipes techniques (par exemple en analyse statistique, en business intelligence, en machine learning ou encore en intelligence artificielle). Livrable : rapport professionnel individuel. Évaluation basée sur la correction du rapport professionnel et une soutenance orale individuelle

Course contents

This course consists of the following modules :

SQL for PostgreSQL

Ref. SGS - 3 days ★ 4 / 5

MongoDB, getting started and development

Ref. MNO - 3 days ★ 4 / 5

Talend Open Studio: Implementing Data Integration

Ref. TOT - 3 days ★ 4 / 5

Web Development, Fundamentals

Ref. DSW - 4 days ★ 3 / 5

Python, developing REST Web Services

Ref. CZT - 2 days

Certification to collect, store and make available data from an artificial intelligence project

Ref. ZDD - 0.5 day

Course schedule

1 RGPD, raising awareness of data protection regulations

- Understand the basic concepts and components of data protection.
- Understand the content of general data protection regulations.
- Know the role of protection authorities.
- Know the legal framework of the RGPD and its scope of application.
- Understand the different stages of compliance and the steps to follow.
- Identify compliance tools.
- Implement a compliance action plan.

TEACHING AIDS AND TECHNICAL RESOURCES

- The main teaching aids and instructional methods used in the training are audiovisual aids, documentation and course material, hands-on application exercises and corrected exercises for practical training courses, case studies and coverage of real cases for training seminars.
- At the end of each course or seminar, ORSYS provides participants with a course evaluation questionnaire that is analysed by our instructional teams.
- A check-in sheet for each half-day of attendance is provided at the end of the training, along with a course completion certificate if the trainee attended the entire session.

TERMS AND DEADLINES

Registration must be completed 24 hours before the start of the training.

ACCESSIBILITY FOR PEOPLE WITH DISABILITIES

Do you need special accessibility accommodations? Contact Mrs. Fosse, Disability Manager, at psh-accueil@orsys.fr to review your request and its feasibility.

2 SQL for PostgreSQL

- Understand the big picture of DBMS.
- Understand the PostgreSQL database.
- Create simple and complex queries.
- Handle internal and external joins.
- Use regular expressions.
- Know the window functions.

3 MongoDB, getting started and development

- Install MongoDB DBMS.
- Configure MongoDB DBMS.
- Handling objects and data in MongoDB.
- Implement a MongoDB application.
- Improving performance.

4 Talend Open Studio, implementing data integration

- Design and develop jobs in Talend's ETL application.
- Optimize jobs by using contexts and datasets.
- Perform more complex transformations using variables, expressions and joins.
- Run and debug a job, trace execution statistics.

5 Developing a website, practical summary

- Understand the fundamentals of the Web.
- Master the technical environment of a website.
- Create a website that is ergonomic, accessible and well referenced.
- Access data in a relational database.
- Manage a website.

6 Web Scraping, harvesting data from the web with Python

- Master the basics of the Python language.
- Advanced programming in Python.
- Get an overview of the main Python libraries available for managing all types of site data.
- Select the right Python library for your web scraping project and be able to implement it.
- Know how to automate large-scale web scraping with scripts.

7 Python Data Science, manipulating and visualizing data

- Get an overview of the Python scientific ecosystem.
- Learn about the essential scientific libraries for data science.
- Be able to manipulate large data sets with Python.
- Understand the benefits of datavisualization.
- Visualize data with Python.

8 Python, developing REST Web Services

- Understand the principles of REST web services.
- Handle JSON data.
- Developing REST APIs with Django REST Framework.
- Securing Web services.

9 **Spark Python, developing applications for big data**

- Discover the fundamental concepts of Spark.
- Use Spark's RDD concept.
- Exploit data with Spark SQL.
- Perform real-time analysis with Spark Streaming.
- Use Spark with Jupyter notebooks, manipulate data with Pyspark as with Pandas.
- Machine learning with Spark;

Dates and locations

REMOTE CLASS
2026 : 2 June, 24 Nov.

PARIS LA DÉFENSE
2026 : 2 June, 24 Nov.