



PROGRAMA FORMATIVO

BIG DATA AWS

Diciembre 2019

DATOS GENERALES DE LA ESPECIALIDAD

1. **Familia Profesional:** INFORMÁTICA Y COMUNICACIONES

Área Profesional: SISTEMAS Y TELEMÁTICA

2. **Denominación:** BIG DATA AWS

3. **Código:** IFCT74

4. **Nivel de cualificación:** 3

5. **Objetivo general:**

Implementar, diseñar y mantener servicios de AWS Big Data de acuerdo con las mejores prácticas de la arquitectura básica, utilizando herramientas para automatizar el análisis de datos proporcionando los conocimientos suficientes para preparar la siguiente certificación oficial de fabricante: AWS Certified Big Data - Specialty

6. **Prescripción de los formadores:**

6.1. Titulación requerida:

Titulación universitaria o Ciclo Formativo de Grado Superior, en su defecto, capacitación profesional equivalente en la especialización relacionada con el curso.

Además de los requisitos anteriores, el docente deberá contar con las certificaciones vigentes y actualizadas en el momento de la impartición de la correspondiente especialidad.

En todo caso, los requisitos mínimos de certificación serán los que marque el fabricante.

6.2. Experiencia profesional requerida:

Al menos doce meses de experiencia profesional en la especialidad objeto, excluyendo la experiencia docente

6.3. Competencia docente:

Será necesario tener experiencia metodológica o experiencia docente contrastada de 500 horas de formación en especialidades relacionadas con la especialidad a impartir.

7. **Criterios de acceso del alumnado:**

7.1. Nivel académico o de conocimientos generales:

- Título de FP Grado Superior o Título de Bachillerato.
- Dominio de inglés a nivel de lectura.
- Conocimientos y/o experiencia profesional en el campo de análisis de datos.
- Conocimiento de algún lenguaje de programación de uso general

Se recomienda disponer de la certificación AWS Certified Developer – Associate y/o AWS Certified Solutions Architect – Associate

Cuando el aspirante al curso no posea el nivel académico indicado, demostrará conocimientos suficientes a través de una prueba de acceso.

8. Número de participantes:

Máximo 25 participantes para cursos presenciales.

9. Relación secuencial de módulos formativos:

- Módulo 1: How to build a Data Lake on AWS.
- Módulo 2: Data Warehouse on AWS.
- Módulo 3: Big Data on AWS.
- Módulo 4: Exam preparation Big Data Specialty.
- Módulo 5: Deep Learning on AWS.

10. Duración:

Horas totales: 120 horas.

Distribución horas:

- Presencial: 120 horas.

11. Requisitos mínimos de espacios, instalaciones y equipamiento

11.1. Espacio formativo:

- Aula de gestión de 3 m² por alumno

Cada espacio estará equipado con mobiliario docente adecuado al número de alumnos, así mismo constará de las instalaciones y equipos de trabajo suficientes para el desarrollo del curso.

11.2. Equipamiento:

- Aula de gestión:
 - Mesa y silla para el formador
 - Mesas y sillas para el alumnado
 - Material de aula
 - Pizarra
 - Impresora láser con conexión a red
 - Pantalla y cañón de proyección
 - PC instalado en red con posibilidad de impresión de documentos, cañón con proyector e internet para el formador (procesador de 64 bits Intel i5 de séptima generación o superior, o AMD similar, 8 GB de memoria R.A.M. o superior, disco duro de 256 GB o superior, preferiblemente SSD, pantalla 17' o superior con resolución mínima WXGA; teclado multimedia USB, ratón sensor óptico USB de 2 botones y rueda de desplazamiento)
 - PCs instalados en red e internet con posibilidad de impresión para los alumnos (procesador de 64 bits Intel i5 de séptima generación o superior, o AMD similar, 8 GB de memoria R.A.M. o superior, disco duro de 256 GB o superior, preferiblemente SSD, pantalla 17' o superior con resolución mínima WXGA, teclado multimedia USB, ratón sensor óptico USB de 2 botones y rueda de desplazamiento)
 - Licencia de sistema operativo
 - Licencia de software antivirus
 - Licencia de software ofimático y herramientas necesarias para la impartición del curso (versión actualizada)
 - Licencias del fabricante para la correcta impartición del curso

En todo caso los requisitos mínimos tanto HW como SW serán los que marque el fabricante como recomendados en cada momento para las versiones actualizadas.

A los alumnos se le proporcionará la documentación oficial de Amazon necesaria para el seguimiento del curso.

Las instalaciones y equipamientos deberán cumplir con la normativa industrial e higiénico sanitaria correspondiente y responderán a medidas de accesibilidad universal y seguridad de los participantes.

En el caso de que la formación se dirija a personas con discapacidad se realizarán las adaptaciones y los ajustes razonables para asegurar su participación en condiciones de igualdad.

12. Requisitos de los centros

Los centros impartidores de formación Oficial de Amazon, deben cumplir el siguiente requisito:

- Acreditación de encontrarse autorizado por AWS como ATP (Authorised Training Partner) de Amazon Web Services.

13. Evaluación del aprendizaje

Se llevará a cabo una evaluación continua y sistemática durante el proceso de aprendizaje y al final del mismo, para comprobar si los alumnos han alcanzado los objetivos establecidos en cada módulo y, por consiguiente, han realizado el curso con el aprovechamiento requerido.

14. Certificación oficial del fabricante

La ejecución y financiación del programa formativo incluye la presentación de los alumnos que han realizado el curso con aprovechamiento a los exámenes para obtener la certificación oficial del fabricante, que gestionará el centro y que en ningún caso supondrá coste alguno para el alumno.

En concreto, para esta acción formativa están incluidos los siguientes exámenes de certificación oficial de Amazon, o los que los sustituyan actualizados al momento de su impartición:

- AWS Certified Big Data - Specialty

MÓDULOS FORMATIVOS

Módulo nº 1

Denominación: HOW TO BUILD A DATA LAKE ON AWS

Objetivo: Diseñar, implementar y operar una solución Data Lake con los servicios de AWS, utilizando las opciones disponibles para el análisis de datos en tiempo 'casi' real.

Duración: 15 horas

Contenidos teórico prácticos:

- Key services that help enable a serverless data lake architecture
- A data analytics solution that follows the ingest, store, process, and analyze workflow
- Repeatable template deployment for implementing a data lake solution
- Building a metadata index and enabling search capability
- Setup of a large scale data ingestion pipeline from multiple data sources
- Transformation of data with simple functions that are event-triggered
- Data processing by choosing the best tools and services for the use case
- Options available to better analyze the processed data
- Best practices for deployment and operations

Módulo nº 2

Denominación: DATA WAREHOUSE ON AWS

Objetivo: Diseñar, implementar y mantener soluciones de almacenamiento de datos basadas en la nube

utilizando Amazon Redshift junto con otros servicios y herramientas de Business Intelligence para la analítica y gestión del dato.

Duración: 40 horas

Contenidos teórico prácticos:

- Introduction to Data Warehousing
- Introduction to Amazon Redshift
- Understanding Amazon Redshift Components and Resources
- Launching an Amazon Redshift Cluster
- Reviewing Data Warehousing Approaches
- Identifying Data Sources and Requirements
- Designing the Data Warehouse
- Loading Data into the Data Warehouse
- Writing Queries and Tuning Performance
- Maintaining the Data Warehouse
- Analyzing and Visualizing Data
- Course Summary

Módulo nº 3

Denominación: BIG DATA ON AWS

Objetivo: Crear entornos de Big Data con los servicios AWS, en base a las mejores prácticas que garantizan un diseño seguro y rentable, utilizando Amazon EMR, junto con las herramientas del ecosistema Hadoop, para el proceso de datos.

Duración: 40 horas

Contenidos teórico prácticos:

- Overview of Big Data
- Big Data Ingestion and Transfer
- Big Data Streaming and Amazon Kinesis
- Lab 1: Using Amazon Kinesis to Stream and Analyze Apache Server Log Data
- Big Data Storage Solutions
- Big Data Processing and Analytics
- Lab 2: Using Amazon Athena to Query Log Data From Amazon S3
- Apache Hadoop and Amazon EMR
- Lab 3: Storing and Querying Data on Amazon DynamoDB
- Using Amazon EMR
- Hadoop Programming Frameworks
- Lab 4: Processing Server Logs With Hive on Amazon EMR
- Web Interfaces on Amazon EMR
- Lab 5: Running Pig Scripts in Hue on Amazon EMR
- Apache Spark on Amazon EMR
- Lab 6: Processing NY Taxi data using Spark on Amazon EMR
- Amazon Redshift and Big Data
- Visualizing and Orchestrating Big Data
- Lab 7: Using TIBCO Spotfire to Visualize Data
- Managing Big Data Costs
- Securing Your Amazon Deployments
- Big Data Design Patterns

Módulo nº 4

Denominación: EXAM PREPARATION BIG DATA SPECIALTY

Objetivo: Superar con éxito el examen de certificación AWS Certified Big Data Specialty.

Duración: 10 horas

Contenidos teórico prácticos:

- Testing center information and expectations
- Exam overview and structure
- Content domains and question breakdown
- Topics and concepts with content domains
- Question structure and interpretation techniques
- Practice exam questions

Módulo nº 5

Denominación: DEEP LEARNING ON AWS

Objetivo: Ejecutar soluciones de Deep Learning en AWS utilizando los servicios Amazon EC2, AMI y el framework MXNet, y explotando el potencial de los servicios AWS Lambda y Amazon EC2 y ECS en su diseño.

Duración: 15 horas

Contenidos teórico prácticos:

- Introduction to Machine Learning
- Introduction to Deep Learning
- Lab 1: Setting up a Deep Learning AMI instance and running a multi-layer perception model
- Introduction to MXNet on AWS
- Lab 2: Running a convolutional neural network model and predicting images on CIFAR-10 dataset
- Deploying Deep Learning Workloads on AWS
- Labs 3: Deploying a Deep Learning model for predicting images using AWS Lambda