

Catálogo de Especialidades Formativas

PROGRAMA FORMATIVO

BIG DATA DEVELOPER CON CLOUDERA APACHE HADOOP

Junio 2025

GOBIERNO DE ESPAÑA
MINISTERIO DE TRABAJO Y ECONOMÍA SOCIAL
SERVICIO PÚBLICO DE EMPLEO ESTATAL
SEPE

SISTEMA NACIONAL DE EMPLEO

IDENTIFICACIÓN DE LA ESPECIALIDAD Y PARÁMETROS DEL CONTEXTO FORMATIVO

Denominación de la especialidad:	BIG DATA DEVELOPER CON CLOUDERA APACHE HADOOP
Familia Profesional:	INFORMÁTICA Y COMUNICACIONES
Área Profesional:	DESARROLLO
Código:	IFCD0170
Nivel de cualificación profesional:	3

Objetivo general

Analizar y abordar casos de estudio utilizando Apache Hadoop y herramientas asociadas al centro de datos empresariales

Relación de módulos de formación

Módulo 1	Desarrollo de aplicaciones Python	75 horas
Módulo 2	Herramientas del ecosistema Hadoop	50 horas
Módulo 3	Desarrollo de aplicaciones en Spark Streaming	50 horas

Modalidades de impartición

Presencial

Duración de la formación

Duración total	175 horas

Requisitos de acceso del alumnado

Acreditaciones / titulaciones	Cumplir como mínimo alguno de los siguientes requisitos: -Título de Bachiller o equivalente -Título de Técnico Superior (FP Grado superior) o equivalente -Haber superado la prueba de acceso a Ciclos Formativos de Grado Superior -Certificado de profesionalidad de nivel 3
Experiencia profesional	Se requiere un mínimo de 6 meses de experiencia desarrollando actividades relacionadas con el manejo de Big Data, programación en Java o Python, o el uso de herramientas del ecosistema Hadoop.
Otros	Cuando no se pueda acreditar ni la titulación ni la experiencia requerida, se deberán demostrar conocimientos sobre: - Programación, preferentemente en Java y/o Python.

Justificación de los requisitos del alumnado
La acreditación de la titulación se justificará mediante la presentación del título o certificado académico correspondiente. La acreditación de la experiencia profesional, mediante la vida laboral. La acreditación de otros requisitos se realizará mediante una prueba o mediante aportación de cualquier evidencia documental de haber cursado una formación en la materia exigida.

Prescripciones de formadores y tutores

Acreditación requerida	Cumplir como mínimo alguno de los siguientes requisitos: • Licenciado/a, Ingeniero/a, Arquitecto/a o el Título de Grado correspondiente u otros títulos equivalentes en los ámbitos de conocimiento de ingeniería informática y de sistemas, ingeniería eléctrica, ingeniería electrónica, ingeniería de la telecomunicación, ingeniería industrial, matemáticas o estadística. • Diplomado/a, Ingeniero/a Técnico/a, Arquitecto/a Técnico/a o el Título de Grado correspondiente y otros títulos equivalentes en los ámbitos de conocmiento de ingeniería informática y de sistemas, ingeniería eléctrica, ingeniería electrónica, ingeniería de la telecomunicación, ingeniería industrial, matemáticas o estadística.
Experiencia profesional mínima requerida	La experiencia profesional requerida será como mínimo de 2 años en el desarrollo de soluciones de Big Data y procesamiento de datos en entornos empresariales, en caso de no cumplir con los requisitos académicos establecidos.
Competencia docente	Se requiere acreditar un mínimo de 500 horas de experiencia docente o 250 horas si la experiencia está relacionada con especialidades o certificados de profesionalidad de la misma familia profesional en la que está encuadrada la presente especialidad.O estar en posesión de alguna de las siguientes titulaciones:-Certificado de Profesionalidad de Docencia de la Formación Profesional para el empleo (SSCE0110) o equivalente.-Certificado de Aptitud Pedagógica (CAP).-Título de licenciado en Pedagogía, Psicopedagogía o de Maestro en cualquiera de sus especialidades, graduado universitario en el ámbito de la Psicología o de la Pedagogía, o un título universitario oficial de posgrado en los citados ámbitos.-Título profesional de Especialización Didáctica, el Certificado de Cualificación Pedagógica, el Máster Universitario habilitante para el ejercicio de las Profesiones Reguladas de Profesor de Educación Secundaria Obligatoria y Bachillerato, Formación Profesional y Escuelas Oficiales de idiomas.

Justificación de las prescripciones de formadores y tutores
La acreditación de la titulación se justificará mediante la presentación del título o certificación académica correspondiente. La experiencia profesional mediante vida

laboral actualizada o certificado de funciones expedido por organización en la que prestó sus servicios. La competencia docente a través de vida laboral actualizada y/o alguna de las titulaciones acreditativas contempladas.

Requisitos mínimos de espacios, instalaciones y equipamientos

Espacios formativos	Superficie m² para 15 participantes	Incremento Superficie/ participante (Máximo 30 participantes)
Aula de informática	45.0 m²	2.4 m² / participante

Espacio formativo	Equipamiento
Aula de informática	-Mesa y silla para el/la formador/a-Mesas y sillas para el alumnado-Material de aula-Pizarra-PC instalado en red con posibilidad de impresión de documentos, cañón con proyector e internet para el personal formador-PCs instalados en red e internet con posibilidad de impresión para el alumnadoTodos los ordenadores han de tener las siguientes características mínimas:- Procesador i5- Memoria RAM 16 GB - Disco duro 500 GB - Tarjeta de red 10/100/1000 Mbps - Tarjeta gráfica- Tarjeta de sonido - Periféricos: Teclado, Ratón y Monitor colorSoftware específico para el aprendizaje de cada acción formativa:- IDEs: IntelliJ, PyCharm o Eclipse -Hadoop y su ecosistema- Apache Hadoop (HDFS, MapReduce).-Apache Spark.-Apache Hive.- Apache Flume.-Apache Sqoop.-Herramientas de gestión de clústeres y Big Data: (Cloudera Manager o Ambari)- Sistema operativo (versión/es con soporte oficial)-Navegador-Software de virtualización: VMware o VirtualBox -Conectividad a Internet: de alta velocidad para descarga de paquetes y acceso a repositorios y plataformas de Big Data.

La superficie de los espacios e instalaciones estarán en función de su tipología y del número de participantes. Tendrán como mínimo los metros cuadrados que se indican para 15 participantes y el equipamiento suficiente para los mismos.

En el caso de que aumente el número de participantes, hasta un máximo de 30, la superficie de las aulas se incrementará proporcionalmente (según se indica en la tabla en lo relativo a m²/participante) y el equipamiento estará en consonancia con dicho aumento.

No debe interpretarse que los diversos espacios formativos identificados deban diferenciarse necesariamente mediante cerramientos.

Las instalaciones y equipamientos deberán cumplir con la normativa industrial e

higiénico-sanitaria correspondiente y responderán a medidas de accesibilidad 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.

Aula virtual

Características
• La impartición de la formación mediante aula virtual se ha de estructurar y organizar de forma que se garantice en todo momento que exista conectividad sincronizada entre las personas formadoras y el alumnado participante así como bidireccionalidad en las comunicaciones. • Se deberá contar con un registro de conexiones generado por la aplicación del aula virtual en que se identifique, para cada acción formativa desarrollada a través de este medio, las personas participantes en el aula, así como sus fechas y tiempos de conexión.
Otras especificaciones
Plataforma de gestión de proyectos o tareas (ej., Trello, Asana) para coordinar trabajos y prácticas de grupo. Acceso a un servidor virtual o clúster configurado para prácticas en tiempo real con herramientas de Big Data (Hadoop, Spark, etc.). VPN o acceso remoto seguro para que los/as participantes puedan conectarse a los recursos del curso de forma segura y sin interrupciones.

Ocupaciones y puestos de trabajo relacionados

- 24161015 ESTADÍSTICOS

- 24161024 TÉCNICOS MEDIOS EN ESTADÍSTICA Y/O CÁLCULO NUMÉRICO

- 27211036 CIENTÍFICOS DE DATOS (BIG DATA)

- 34041019 TÉCNICOS EN ESTADÍSTICA

- 38201017 PROGRAMADORES DE APLICACIONES INFORMÁTICAS

Requisitos oficiales de las entidades o centros de formación

Estar inscrito en el Registro de entidades de formación (Servicios Públicos de Empleo).

DESARROLLO MODULAR

MÓDULO DE FORMACIÓN 1: Desarrollo de aplicaciones Python

OBJETIVO

Desarrollar aplicaciones en Python.

DURACIÓN:

75 horas

RESULTADOS DE APRENDIZAJE

Conocimientos / Capacidades cognitivas y prácticas

- Adquisición de fundamentos de programación en Python
    - Estructuras de datos en Python (listas, diccionarios, tuplas)
    - Principios de programación orientada a objetos
    - Uso de bibliotecas y módulos en Python
    - Introducción a la gestión de excepciones y manejo de errores
- Comprensión de la sintaxis y semántica de Python
    - Capacidad para analizar problemas y descomponerlos en pasos lógicos
    - Evaluación de soluciones de programación para optimizar el rendimiento
    - Interpretación y modificación de código existente
- Desarrollo de scripts y programas en Python
    - Implementación de estructuras de control (condicionales, bucles)
    - Creación y manipulación de funciones y módulos
    - Actividades que impliquen el uso de Python para resolver problemas específicos
- Aplicación de medidas de eficiencia energética y sostenibilidad ambiental - Prácticas de desarrollo sostenible (green coding) -Estrategias para reducir la huella de carbono en el desarrollo y despliegue de aplicaciones

Habilidades de gestión, personales y sociales

- Valoración de la importancia del cumplimiento de plazos en el desarrollo de proyectos de programación.
- Comprensión de las necesidades y perspectivas de los compañeros de equipo.
- Aplicación de medidas de eficiencia energética y sostenibilidad ambiental en el desarrollo de software.

MÓDULO DE FORMACIÓN 2: Herramientas del ecosistema Hadoop

OBJETIVO

Importar datos a Apache Hadoop Cluster y procesarlos con Spark, Hive, Flume, Sqoop, Impala, y otras herramientas del ecosistema de Hadoop.

DURACIÓN:

50 horas

RESULTADOS DE APRENDIZAJE

Conocimientos / Capacidades cognitivas y prácticas

- Conocimiento del ecosistema de Hadoop y su arquitectura
    - Conceptos básicos de Apache Spark y su funcionamiento
    - Uso de RDDs (Resilient Distributed Datasets)
    - Importación de datos utilizando Apache Sqoop
    - Introducción a Hive y la gestión de datos con Impala
- Identificación y selección de la herramienta adecuada para cada tarea dentro del ecosistema Hadoop
    - Evaluación de modelos de datos y su optimización en Spark
    - Análisis de datos y transformación utilizando Spark
    - Desarrollo de soluciones escalables para el procesamiento de datos
- Realización de operaciones con herramientas del ecosistema Hadoop
    - Configuración y uso de clústeres Hadoop
    - Desarrollo de aplicaciones Spark para el procesamiento de datos
    - Carga y transformación de datos con Sqoop
    - Realización de consultas y análisis de datos con Hive

Habilidades de gestión, personales y sociales

- Adaptación a las tecnologías del ecosistema Hadoop.
- Fomento de actitudes positivas ante la colaboración a través de equipos de trabajo para implementar soluciones de Big Data.

MÓDULO DE FORMACIÓN 3: Desarrollo de aplicaciones en Spark Streaming

OBJETIVO

Simplificar el desarrollo con Kite SDK, definir y usar Data Sets, importar datos relacionales con Apache Sqoop, capturar datos con ApacheFlume y desarrollar componentes personalizados.

DURACIÓN:

50 horas

RESULTADOS DE APRENDIZAJE

Conocimientos / Capacidades cognitivas y prácticas

- Desarrollo de aplicaciones en Spark Streaming
    - Uso de herramientas de captura de datos como Apache Flume
    - Gestión de workflows con Apache Oozie.Implementación de componentes personalizados y transformación de datos
    - Técnicas de análisis de datos en tiempo real
- Diseño de flujos de trabajo complejos en entornos distribuidos
    - Evaluación de estrategias de procesamiento de datos en tiempo real
    - Solución de problemas y optimización de aplicaciones Spark
    - Aplicación de técnicas de Machine Learning en el contexto de Big Data
- Desarrollo y despliegue de aplicaciones Spark Streaming
    - Implementación de pipelines de datos con Apache Flume
    - Creación y gestión de workflows usando Oozie
    - Análisis de datos en tiempo real, incorporando datos de diferentes fuentes y formatos

Habilidades de gestión, personales y sociales

- Desarrollo de hablidades para la comunicación de resultados y análisis a diferentes audiencias.
- Fomento de un ambiente de colaboración en el contexto de equipos de trabajo, valorando las ideas y aportes de los demás.

EVALUACIÓN DEL APRENDIZAJE EN LA ACCIÓN FORMATIVA

La evaluación tendrá un carácter teórico-práctico y se realizará de forma sistemática y continua, durante el desarrollo de cada módulo y al final del curso.

Puede incluir una evaluación inicial de carácter diagnóstico para detectar el nivel de partida del alumnado.

La evaluación se llevará a cabo mediante los métodos e instrumentos más adecuados para comprobar los distintos resultados de aprendizaje, y que garanticen la fiabilidad y validez de la misma.

Cada instrumento de evaluación se acompañará de su correspondiente sistema de corrección y puntuación en el que se explicite, de forma clara e inequívoca, los criterios de medida para evaluar los resultados alcanzados por los participantes.

La puntuación final alcanzada se expresará en términos de Apto/ No Apto.