

# Big Data, síntesis técnica

Curso de síntesis de 2 días - 14h

Ref.: BAG - Precio 2024: 1 360€ sin IVA

En esta síntesis se presentarán los retos y las aportaciones de los macrodatos o Big Data, así como las tecnologías disponibles para su implantación. Seguirá los pasos de un proyecto de datos masivos desde la instalación de una plataforma de Big Data, la ingestión y el procesamiento de datos hasta la visualización de los resultados.

## OBJETIVOS PEDAGÓGICOS

Al término de la formación, el alumno podrá:

Descubrir los conceptos clave de los Big Data

Comprender el ecosistema tecnológico de un proyecto de Big Data

Evaluar las técnicas de gestión de flujos de datos masivos

Implementar modelos de análisis estadísticos para responder a las necesidades de la empresa

Descubrir las herramientas de visualización de datos

Presentar la plataforma Hadoop y sus componentes básicos, utilizar un ETL para gestionar los datos, crear modelos de análisis y cuadros de control.

## PROGRAMA

última actualización: 01/2023

### 1) Comprender los conceptos clave y los retos de los Big Data

- Los orígenes de los Big Data.
- El valor de los datos: un cambio importante.
- Los datos como materia prima.
- Las cifras clave del mercado en el mundo y en Francia.
- Los retos de los Big Data: Rendimiento de la inversión (ROI), organización y confidencialidad de los datos.

*Demostración* : Presentación de una arquitectura de Big Data.

### 2) Tecnologías de Big Data

- Arquitectura y componentes de la plataforma Hadoop 2.
- Modos de almacenamiento (NoSQL, HDFS).
- Funcionamiento de MapReduce y Yarn...
- Principales distribuciones de Hadoop: Hortonworks, Cloudera, MapR, etc.
- Tecnologías emergentes: Spark, Storm, Machine Learning Azure, etc.
- Cómo instalar una plataforma Hadoop.
- Presentación de tecnologías específicas para Big Data (Talend, Tableau, Qlikview, etc.).

*Demostración* : Instalación de una plataforma completa de Big Data.

### 3) Procesamiento de Big Data

- Funcionamiento del sistema de archivos distribuido Hadoop (HDFS por su sigla en inglés).
- Importar datos a HDFS.
- Tratamiento de datos con PIG.
- Consultas SQL con HIVE.

## PARTICIPANTES

Minería de datos, investigadores estadísticos, desarrolladores, gestores de proyecto y consultores de inteligencia empresarial.

## REQUISITOS PREVIOS

Conocimientos básicos de modelos relacionales, estadística y lenguajes de programación. Conocimientos básicos de los conceptos de la Inteligencia Empresarial (Business Intelligence).

## COMPETENCIAS DEL FORMADOR

Los expertos que imparten la formación son especialistas en las materias tratadas. Han sido validados por nuestros equipos pedagógicos, tanto en el plano de los conocimientos profesionales como en el de la pedagogía, para cada curso que imparten. Cuentan al menos con entre cinco y diez años de experiencia en su área y ocupan o han ocupado puestos de responsabilidad en empresas.

## MODALIDADES DE EVALUACIÓN

El formador evalúa los progresos pedagógicos del participante a lo largo de toda la formación mediante preguntas de opción múltiple, escenificaciones de situaciones, trabajos prácticos, etc.

El participante también completará una prueba de posicionamiento previo y posterior para validar las competencias adquiridas.

## MEDIOS PEDAGÓGICOS Y TÉCNICOS

- Los medios pedagógicos y los métodos de enseñanza utilizados son principalmente: ayudas audiovisuales, documentación y soporte de cursos, ejercicios prácticos de aplicación y ejercicios corregidos para los cursillos prácticos, estudios de casos o presentación de casos reales para los seminarios de formación.
- Al final de cada cursillo o seminario, ORSYS facilita a los participantes un cuestionario de evaluación del curso que analizarán luego nuestros equipos pedagógicos.
- Al final de la formación se entrega una hoja de presencia por cada media jornada de presencia, así como un certificado de fin de formación si el alumno ha asistido a la totalidad de la sesión.

## MODALIDADES Y PLAZOS DE ACCESO

La inscripción debe estar finalizada 24 horas antes del inicio de la formación.

## ACCESIBILIDAD DE LAS PERSONAS CON DISCAPACIDAD

¿Tiene alguna necesidad específica de accesibilidad? Póngase en contacto con la Sra. FOSSE, interlocutora sobre discapacidad, en la siguiente dirección psh-accueil@orsys.fr para estudiar de la mejor forma posible su solicitud y su viabilidad.

- Creación de flujos de datos masivos con un ETL.

*Demostración : Implementación de flujos de datos masivos.*

#### 4) Métodos de análisis y tratamiento de datos para Big Data

- Métodos de exploración.
- Segmentación y clasificación.
- Aprendizaje automático, estimación y predicción.
- Tiempo real e IA.
- Aplicación de modelos.

*Demostración : Presentación del entorno Spark, Jupyter Notebook, R Notebook y Shiny.*

*Configuración de análisis de aprendizaje automáticos con los lenguajes R, Python y Scala.*

#### 5) Visualización de datos y representación de los datos de forma visual

- Principales soluciones en el mercado.
- Ir más allá de los informes estáticos.
- La visualización de datos y el arte de contar las cifras de forma creativa y lúdica.
- Medir la reputación online, la notoriedad de una marca, la experiencia y la satisfacción de los clientes...

*Demostración : Presentación y utilización de una herramienta de visualización de datos para crear análisis dinámicos.*

#### 6) Conclusión

- Las condiciones del éxito.
- Síntesis de las mejores prácticas.
- Bibliografía.

## FECHAS

---

Contacto