

The background of the page is a blurred image of a person's hand using a computer mouse. On the right side, there is a vertical column of seven white circles of varying sizes, with the largest one at the bottom. The text is overlaid on a semi-transparent grey rounded rectangle.

## **IFCT159. Introducción al Big Data e Inteligencia Artificial**

## Objetivos

---

### □ **Objetivo General**

- Conocer las principales características del procesamiento masivo de datos en su relación con aplicaciones y algoritmos de Inteligencia Artificial, así como facilitar una aproximación a la planificación y desarrollo de programas de inteligencia artificial y Big Data en el entorno empresarial.

### □ **Objetivos Específicos**

- Comprender la evolución de la inteligencia empresarial tradicional hacia el Big Data, analizando casos prácticos como navegación web, geolocalización y audiencias de televisión.
- Identificar las aplicaciones empresariales clave del Big Data y su impacto en sectores como la salud, el comercio y las telecomunicaciones.
- Explorar el funcionamiento y la utilidad de herramientas tecnológicas esenciales como Hadoop, HDFS y MapReduce para el tratamiento masivo de datos.
- Analizar las características fundamentales del Big Data, incluyendo volumetría, velocidad, variedad, veracidad y valor del dato, mediante ejemplos prácticos.
- Evaluar los beneficios y desafíos de los procesos en tiempo real y del cloud computing como habilitadores claves del Big Data en la toma de decisiones estratégicas.
- Comprender los fundamentos de la arquitectura Big Data, incluyendo el ecosistema Hadoop y sus componentes principales.
- Identificar los lenguajes de programación más utilizados en el análisis y procesamiento de datos masivos, así como sus aplicaciones específicas.
- Explorar las herramientas clave de procesos ETL y sistemas de datos en tiempo real para asegurar una gestión eficiente de la información.
- Reconocer la importancia de la seguridad y el gobierno del dato como pilares fundamentales para la integridad y el uso responsable de la información.
- Comprender los conceptos básicos de la Ciencia de Datos y la Inteligencia Artificial, incluyendo su relación con el Big Data.
- Identificar y utilizar los lenguajes de programación más relevantes, como R y Python, con un enfoque en sus aplicaciones en IA y Ciencia de Datos.

- Analizar y aplicar algoritmos supervisados y no supervisados en escenarios prácticos, considerando sus beneficios y limitaciones.
- Explorar herramientas y técnicas para el procesamiento de información no estructurada y la visualización de datos, desarrollando habilidades para interpretar resultados de forma efectiva.
- Comprender el impacto del Big Data en las instituciones públicas y su relación con el concepto de OpenData.
- Analizar la aplicabilidad del Big Data en el ámbito empresarial para mejorar la eficiencia operativa y la toma de decisiones estratégicas.
- Explorar el uso del Big Data en proyectos con impacto social, dentro de la iniciativa "Data for Good".
- Identificar las tendencias futuras del Big Data, incluyendo la automatización, la personalización y los desafíos éticos y legales asociados.

## Contenidos

<b>IFCT159. Introducción al Big Data e inteligencia artificial</b>	<b>Tiempo estimado</b>
<p><b>Unidad 1:</b> Adquisición y dominio de conceptos básicos y de conocimientos sobre los avances en Big Data.</p> <ol style="list-style-type: none"> <li>1. Cómo evoluciona el BI tradicional al Big Data (navegación web, geolocalización, audiencias tv, etc.)               <ol style="list-style-type: none"> <li>1.1. Ejemplos y casos representativos (navegación web, geolocalización, audiencias tv, etc.)</li> </ol> </li> <li>2. El Big Data como solución al tratamiento masivo de datos: definición e historia desde su invención               <ol style="list-style-type: none"> <li>2.1. Aplicaciones empresariales clave: sectores y áreas donde Big Data ha marcado la diferencia</li> </ol> </li> <li>3. Hadoop como revolución para el tratamiento paralelo de datos masivo               <ol style="list-style-type: none"> <li>3.1. HDFS y MapReduce</li> </ol> </li> <li>4. Características del Big Data (4Vs y más): volumetría, velocidad, variedad, veracidad (calidad del dato), valor del dato, etc.               <ol style="list-style-type: none"> <li>4.1. Volumetría, velocidad, variedad y veracidad: ejemplos prácticos</li> <li>4.2. El valor del dato como ventaja competitiva</li> </ol> </li> <li>5. Nuevos paradigmas del Big Data: procesos en real time y cloud computing               <ol style="list-style-type: none"> <li>5.1. Procesos en real time: entender los macrodatos en tiempo real</li> <li>5.2. Cloud computing como habilitador del Big Data: beneficios, desafíos y ejemplos de su implementación</li> </ol> </li> </ol>	
Cuestionario de autoevaluación Unidad 1	<b>30 minutos</b>
Actividad de evaluación Unidad 1	<b>90 minutos</b>
Tiempo total de la unidad	<b>9,30 horas</b>

**Unidad 2:** Conocimiento de nociones básicas sobre arquitectura Big Data y principales tecnologías.

1. El ecosistema Hadoop: HDFS y MapReduce
  - 1.1. HDFS: fundamentos y almacenamiento distribuido
  - 1.2. MapReduce: modelo de procesamiento paralelo
2. Principales lenguajes de programación utilizados: Java, Scala, SQL y Python
  - 2.1. Java y Scala: principales usos y casos de ejemplo
  - 2.2. SQL y Python: principales usos y casos de ejemplo
3. Procesos ETL (Extracción, Transformación y Carga): Flume, Sqoop y Hive
  - 3.1. El uso de Flume y Sqoop como herramientas principales
  - 3.2. Hive como herramienta para el análisis de datos
4. Procesos real time y bases de datos de alta disponibilidad: Kafka, HBase y Redis
  - 4.1. Kafka: gestión de datos en tiempo real
  - 4.2. HBase y Redis: bases de datos distribuidas y de alta disponibilidad
5. Procesamiento y analítica avanzada con Spark
6. Seguridad y gobierno del dato

Cuestionario de autoevaluación Unidad 2	<b>30 minutos</b>
Actividad de evaluación Unidad 2	<b>90 minutos</b>
Tiempo total de la unidad	<b>10 horas</b>

**Unidad 3:** Comprensión de los principales conceptos sobre la Ciencia de Datos e IA.

1. Introducción a la Ciencia de Datos y la Inteligencia Artificial
  - 1.1. Principales diferencias entre Ciencia de Datos y Big Data
2. Principales lenguajes de programación utilizados: R y Python
  - 2.1. Python: librerías específicas para IA y Ciencia de Datos
3. Algoritmos supervisados: ¿qué son? algunos ejemplos
  - 3.1. Regresión lineal y árboles de decisión
4. Algoritmos no supervisados: ¿qué son? algunos ejemplos
5. Introducción al Deep Learning y al Aprendizaje por Refuerzo
  - 5.1. Redes neuronales profundas: conceptos básicos
6. Procesamiento de información no estructurada: imágenes y textos
  - 6.1. Herramientas para el procesamiento de información no estructurada
7. Visualización de datos: visualizaciones interactivas y dashboards

Cuestionario de autoevaluación Unidad 3	<b>30 minutos</b>
Actividad de evaluación Unidad 3	<b>90 minutos</b>
Tiempo total de la unidad	<b>10 horas</b>

<p><b>Unidad 4:</b> Adquisición de una visión transversal sobre el futuro del Big Data y cómo se aplica actualmente en diferentes áreas.</p> <ol style="list-style-type: none"> <li>1. Ejemplos en las instituciones públicas: OpenData             <ol style="list-style-type: none"> <li>1.1. Iniciativas de OpenData a nivel internacional</li> <li>1.2. Impacto del OpenData en la transparencia gubernamental</li> </ol> </li> <li>2. Ejemplos en el mundo empresarial: ejemplos de aplicabilidad del Big Data a la eficiencia de las operaciones de una compañía             <ol style="list-style-type: none"> <li>2.1. Aplicaciones del Big Data en la optimización de operaciones</li> <li>2.2. Big Data en la toma de decisiones estratégicas</li> </ol> </li> <li>3. Data For Good: Big Data para el bien social             <ol style="list-style-type: none"> <li>3.1. Ejemplos de proyectos sociales impulsados por el Big Data</li> </ol> </li> <li>4. Reflexiones finales sobre el impacto del Big Data en los años venideros             <ol style="list-style-type: none"> <li>4.1. Principales tendencias futuras: automatización y personalización</li> <li>4.2. Retos éticos y legales en la evolución del Big Data</li> </ol> </li> </ol>	
Cuestionario de autoevaluación Unidad 4	<b>30 minutos</b>
Actividad de evaluación Unidad 4	<b>90 minutos</b>
Tiempo total de la unidad	<b>9,30 horas</b>
Examen final	<b>1 hora</b>
<b>4 unidades</b>	<b>40 horas</b>