



### 1. Datos Generales de la asignatura.

<b>Nombre de la asignatura:</b>	<b>Ingeniería de los Datos.</b>
<b>Clave de la asignatura:</b>	<b>ADD-2304.</b>
<b>SATCA<sup>1</sup>:</b>	<b>2-3-5</b>
<b>Carrera:</b>	<b>Ingeniería en Sistemas Computacionales.</b>

### 2. Presentación.

<p><b>Caracterización de la asignatura.</b></p> <p><i>Esta asignatura aporta al perfil del estudiante los conocimientos necesarios para el desarrollo, implementación, administración, el almacenamiento y la gestión de datos en una organización con el objetivo de detectar tendencias, y el proceso de los mismos de forma más útil.</i></p> <p><i>Para cursar esta asignatura se requiere tener habilidades de programación y administración de bases de datos relacionales y no-relacionales, así como de cómputo en la nube; es necesario contar con las competencias desarrolladas en las materias previas de: Fundamentos de Programación, Fundamentos de Base de Datos, Taller de Bases de Datos, Administración de Bases de Datos y Cómputo en la Nube en temas manejo de una base de datos.</i></p> <p><i>Contribuye activamente en la generación de conocimientos a través de las experiencias de aprendizaje que se desarrollen a lo largo del curso, acoplándose a las líneas de investigación definidas por la carrera de Ingeniería en Sistemas Computacionales.</i></p>
<p><b>Intención didáctica.</b></p> <p>Esta asignatura está organizada en seis temas. Se inicia el curso con el tema 1 en el cual se analiza la perspectiva del Big Data el cual se ocupa de la procedencia y la naturaleza de las grandes cantidades de información que procesa y el Machine Learning mediante análisis inteligente se encarga del aprendizaje computacional.</p> <p>En el tema 2, se aborda el Análisis, Adquisición y el Dashboard, que incluye el procesamiento de los datos, creación de web scraping (técnica en la cual un programa informático extrae datos del resultado generado por otro programa) y mostrando los datos mediante un dashboard</p>

<sup>1</sup> Sistema de Asignación y Transferencia de Créditos Académicos





(visualización de datos en forma de gráficas para que los usuarios puedan hacer un seguimiento de la salud de su negocio en relación con los puntos de referencia y los objetivos.

El tema 3, se enfoca a la extracción y conexión de datos mediante tecnologías json y api, tomando como referencia bases de datos NoSql.

En el tema 4, se aborda el tema Data Warehouse cuyo objetivo es almacenar datos de una organización que facilite la toma de decisiones en la administración. No solo incluye datos, sino herramientas, procedimientos que facilitan el acceso a los datos para aquellos que toman decisiones.

En el tema 5, se abordan herramientas de Spark la cual permite que los programadores escriban aplicaciones rápidamente en Java, Scala, Python, R y SQL para que puedan acceder a ellas los desarrolladores, los científicos de datos y los usuarios avanzados, así como el Data lake que es un repositorio de almacenamiento que contienen una gran cantidad de datos en bruto y que se mantienen allí hasta que sea necesario.

Por último, en el tema 6 se enfoca a la automatización de Data Pipelines el cual es un servicio web que se puede utilizar para automatizar el movimiento y la transformación de los datos. Se pueden definir flujos de trabajo controlados por datos en los que las tareas dependen de que se hayan realizado correctamente las tareas anteriores.

Se sugiere desarrollar experiencias de aprendizaje basado en proyectos que permitan la integración de los contenidos y contribuyan al desarrollo de competencias instrumentales, conceptuales y actitudinales de los estudiantes.; así mismo se recomienda la realización de un proyecto integrador con las materias del módulo especialidad que incorporen competencias específicas.

Los contenidos se abordarán de manera secuencial como los marca el programa, buscando la aplicación del conocimiento en el proyecto integrador de la asignatura, incorporando de manera progresiva los temas revisados con un enfoque basado en actividades que promuevan en el estudiante el desarrollo de sus habilidades para trabajar en equipo y aplicar el conocimiento a la práctica.

El docente además de ser un motivador permanente en el proceso educativo deberá ser promotor y director de la enseñanza a través de la transmisión de su conocimiento, así como la aplicación de sus habilidades y destrezas utilizando las herramientas tradicionales y digitales a su alcance para cautivar a sus estudiantes e interesarlos en el tema.

### 3. Participantes en la actualización, el diseño, consolidación y/o seguimiento curricular del programa.



Número de registro: 89HL-072  
Fecha de inicio: 2017-04-10  
Término de la certificación 2021-04-10

Av. Tecnológico S/N, Col. La Primavera, C.P. 47829 Ocotlán, Jalisco. Tel. 01 (392) 92 24630 y 92 53057, Ext. 120 e-mail: [aurora.nn@ocotlan.tecnm.mx](mailto:aurora.nn@ocotlan.tecnm.mx) [tecnm.mx](http://tecnm.mx) | [itocotlan.com](http://itocotlan.com)



**2022** Ricardo Flores  
Año de Magón  
PRECURSOR DE LA REVOLUCIÓN MEXICANA





Lugar y fecha de elaboración o revisión.	Participantes.	Observaciones.
<i>Instituto Tecnológico de Ocotlán Agosto 2022.</i>	<i>Revisión y aprobación por: Integrantes de la Academia de Sistemas y Computación. Elaboración: Aurora Berenice Navarro Nuñez. María Elizabeth Bautista Padilla. Jose Luis Castellanos Gutiérrez. Magda Sagrario Velázquez Lopez.</i>	<i>Instituto Tecnológico de Ocotlán.  Definición de los programas de estudio de especialidad para la carrera de Ingeniería en Sistemas Computacionales. ISIC-2010-224.</i>

#### 4. Competencia(s) a desarrollar.

##### Competencia(s) específica(s) de la asignatura.

Analiza, diseña e implementa soluciones para el almacenamiento de datos, basado en las tecnologías en la nube y desarrollando algoritmos que permitan procesar información para la visualización, análisis de negocios y soluciones computacionales en una organización.

#### 5. Competencias previas.

- Diseño, implementación y administración de Bases de Datos.
- Analizar y diseñar soluciones de aplicaciones informáticas para ofrecer servicios de computación en la nube como un nuevo modelo de prestación de servicios de negocio y tecnologías.
- Desarrolla aplicaciones basadas en lenguajes y herramientas de código libre, utilizados en la web y dispositivos móviles.

#### 6. TEMARIO.



Av. Tecnológico S/N, Col. La Primavera, C.P. 47829 Ocotlán, Jalisco. Tel. 01 (392) 92 24630 y 92 53057, Ext. 120 e-mail: [aurora.nn@ocotlan.tecnm.mx](mailto:aurora.nn@ocotlan.tecnm.mx) [tecnm.mx](http://tecnm.mx) | [itocotlan.com](http://itocotlan.com)



No.	Temas.	Subtemas.
1	Fundamentos del Big Data y Machine Learning.	1.1 Introducción al Big Data y Machine Learning. 1.2.1 Importancia, características y campos de aplicación. 1.4 Modelo y Componentes. 1.5 Herramientas y Tecnología. 1.6 Equipos de alto rendimiento: su integración y desempeño.
2	Análisis, adquisición y Dashboard.	2.1 Análisis. 2.1.1 Análisis de Big Data. 2.1.2 Análisis de sentimientos. 2.1.3 Machine Learning.  2.2 Adquisición. 2.2.1 Web Scraping. 2.2.2 Adquisición de datos de páginas Web.  2.3 Dashboard. 2.3.1 Estructura de un dashboard. 2.3.2 Creación de gráficos en tiempo real y en web. 2.3.4 Creación de un servicio web. 2.3.5 Muestra los datos en web.
3	Extracción y Conexión.	3.1 Extracción de Datos. 3.1.1 JSON. 3.1.2 API. 3.1.3 Extracción de datos en redes sociales. 3.1.5 Extracción de archivos de audio y video.  3.2 Conexión de Datos. 3.2.1 Conexión a base de datos NoSQL. 3.2.2 Manejo de bases de datos No SQL. 3.2.3 Repositorio de datos. 3.3 Liderazgo y coaching: sus diferencias y su aportación en el trabajo en equipo.



4	Data Warehouses.	4.1 Introducción al Data Warehouses. 4.2 Construcción de una Data Warehouse. 4.3 Herramientas y tecnología en la nube para Data Warehouse. 4.4 Desarrollo de un Networking en la gestión de tecnología.
5	Spark y Data Lakes.	5.1 Construcción de un Data Lake. 5.2 Uso del Spark. 5.2.1 Data Wrangling con Spark. 5.3 Depuración y optimización.
6	Automatización de Data Pipelines.	6.1 Data pipelines con airflow. 6.1.1 Data pipelines. 6.1.2 Data Quality. 6.1.3 Generación de Data Pipelines.
7	Proyecto final.	7.1 Implementación. 7.2 Justificación. 7.3 Documentación de soporte. 7.4 Presentación de alto impacto Blitz.

## 7. Actividades de aprendizaje de los temas.

<b>Fundamentos del Big Data y Machine Learning.</b>	
Competencias.	Actividades de aprendizaje.
<p><b>Específica(s):</b> Analizar la perspectiva del Big Data el cual se ocupa de la procedencia y la naturaleza de las grandes cantidades de información que procesa y el Machine Learning mediante análisis inteligente que se encarga del aprendizaje computacional.</p> <p><b>Genéricas:</b></p> <ul style="list-style-type: none"> <li>● Capacidad de análisis y síntesis.</li> <li>● Capacidad de aplicar los conocimientos en la práctica.</li> <li>● Capacidad de organizar y planificar.</li> <li>● Habilidad para buscar y analizar información proveniente de fuentes diversas.</li> <li>● Habilidades de investigación.</li> </ul>	<ul style="list-style-type: none"> <li>● Investigar los temas de Big Data, Machine Learning, importancia, características y campos de aplicación.</li> <li>● Investigar las herramientas de software actuales con las cuales es viable aplicar estas tecnologías.</li> <li>● Realizar ejercicios para la implementación de big data y machine learning.</li> </ul>

<ul style="list-style-type: none"> <li>● Capacidad de generar nuevas ideas.</li> <li>● Toma de decisiones.</li> <li>● Trabajo en equipo.</li> <li>● Liderazgo.</li> <li>● Habilidad para trabajar en forma. Autónoma.</li> <li>● Búsqueda del logro.</li> </ul>	
<b>Análisis, Adquisición y Dashboard.</b>	
Competencias.	Actividades de aprendizaje.
<p>Específica(s): Procesar los datos mediante la creación de web scraping (técnica en la cual un programa informático extrae datos del resultado generado por otro programa) y mostrar los datos mediante un dashboard (visualización de datos en forma de gráficas para que los usuarios puedan hacer un seguimiento de la salud de su negocio en relación con los puntos de referencia y los objetivos.</p> <p>Genéricas:</p> <ul style="list-style-type: none"> <li>● Capacidad de análisis y síntesis.</li> <li>● Capacidad de pensamiento divergente.</li> <li>● Capacidad de aplicar los conocimientos en la práctica.</li> <li>● Capacidad de organizar y planificar.</li> <li>● Habilidad para buscar y analizar información proveniente de fuentes diversas.</li> <li>● Habilidades de investigación.</li> <li>● Capacidad de generar nuevas ideas.</li> <li>● Toma de decisiones.</li> <li>● Trabajo en equipo.</li> <li>● Liderazgo.</li> <li>● Habilidad para trabajar en forma. Autónoma.</li> <li>● Búsqueda del logro.</li> </ul>	<ul style="list-style-type: none"> <li>● Investigar técnicas para el análisis de big data, métodos para aplicar el análisis de sentimientos e implementación del machine learning.</li> <li>● Realizar ejercicios diseñados para el análisis de big data y análisis de sentimientos.</li> <li>● Programar elementos para la aplicación de análisis de big data y análisis de sentimientos.</li> <li>● Investigar técnicas para la adquisición de datos mediante la aplicación de web scraping.</li> <li>● Realizar ejercicios para la adquisición de datos de páginas web.</li> <li>● Investigar conceptos de estructuras de un dashboard.</li> <li>● Realizar ejercicios para la creación de gráficas en tiempo real y en la web mediante un dashboard.</li> </ul>
<b>Extracción y Conexión.</b>	
Competencias.	Actividades de aprendizaje.
Específica(s):	



<p>Extraer y conectar los datos mediante tecnologías json y api, tomando como referencia bases de datos NoSql.</p> <p>Genéricas:</p> <ul style="list-style-type: none"> <li>● Capacidad de análisis y síntesis</li> <li>● Capacidad de aplicar los conocimientos en la práctica.</li> <li>● Capacidad de organizar y planificar.</li> <li>● Habilidad para buscar y analizar información proveniente de fuentes diversas.</li> <li>● Habilidades de investigación.</li> <li>● Capacidad de generar nuevas ideas.</li> <li>● Capacidad de generar redes de contacto profesionales.</li> <li>● Toma de decisiones.</li> <li>● Trabajo en equipo.</li> <li>● Liderazgo.</li> <li>● Habilidad para trabajar en forma. Autónoma.</li> <li>● Búsqueda del logro.</li> </ul>	<ul style="list-style-type: none"> <li>● Investigar los temas y conceptos claves sobre extracción de datos mediante JSON, API así como de redes sociales en formatos de audio y video.</li> <li>● Investigar los temas y conceptos claves sobre conexión de datos mediante en bases de datos nosql y repositorio de datos.</li> <li>● Realizar ejercicios prediseñados y propuestos en la documentación oficial.</li> <li>● Programar los elementos de un proyecto que el docente considere conveniente, orientado a la introducción de datawarehouse y minería de datos.</li> </ul>
<b>Data Warehouses.</b>	
Competencias.	Actividades de aprendizaje.
<p>Específica(s): Almacenar datos de una organización que facilite la toma de decisiones en la administración utilizando herramientas, procedimientos que facilitan el acceso a los datos para aquellos que toman decisiones.</p> <p>Genéricas:</p> <ul style="list-style-type: none"> <li>● Capacidad de análisis y síntesis.</li> <li>● Capacidad de aplicar los conocimientos en la práctica.</li> <li>● Capacidad de organizar y planificar.</li> <li>● Habilidad para buscar y analizar información proveniente de fuentes diversas.</li> <li>● Habilidades de investigación.</li> <li>● Capacidad de generar nuevas ideas.</li> <li>● Toma de decisiones.</li> <li>● Trabajo en equipo.</li> </ul>	<ul style="list-style-type: none"> <li>● Investigar los temas y conceptos claves sobre data warehouse, su construcción y las herramientas en las cuales se pueden implementar.</li> <li>● Realizar ejercicios que permitan realizar aplicaciones de data warehouse en la nube.</li> <li>● Programar los elementos de un proyecto que el docente considere conveniente, de tal forma que el estudiante pueda hacer la aplicación de data warehouse.</li> </ul>



<ul style="list-style-type: none"> <li>• Liderazgo.</li> <li>• Habilidad para trabajar en forma. Autónoma.</li> <li>• Búsqueda del logro.</li> </ul>	
--	--

<b>Spark y Data Lakes.</b>	
Competencias.	Actividades de aprendizaje.
<p><b>Específica(s):</b> Escribir aplicaciones rápidamente en Java, Scala, Python, R y SQL para que puedan acceder a ellas los desarrolladores, los científicos de datos y los usuarios avanzados, así como el Data lake que es un repositorio de almacenamiento que contienen una gran cantidad de datos en bruto y que se mantienen allí hasta que sea necesario.</p> <p><b>Genéricas:</b></p> <ul style="list-style-type: none"> <li>• Capacidad de análisis y síntesis.</li> <li>• Capacidad de aplicar los conocimientos en la práctica.</li> <li>• Capacidad de organizar y planificar.</li> <li>• Habilidad para buscar y analizar información proveniente de fuentes diversas.</li> <li>• Habilidades de investigación.</li> <li>• Capacidad de generar nuevas ideas.</li> <li>• Toma de decisiones.</li> <li>• Trabajo en equipo.</li> <li>• Liderazgo.</li> <li>• Habilidad para trabajar en forma. Autónoma.</li> <li>• Búsqueda del logro.</li> </ul>	<ul style="list-style-type: none"> <li>• Investigar los temas y conceptos claves sobre la construcción de data lakes, spark y la creación de data wrangling.</li> <li>• Realizar ejercicios aplicando en los data lakes la depuración y optimización.</li> <li>• Programar los elementos de un proyecto que el docente considere conveniente, de tal forma que el estudiante pueda hacer implementaciones de data lakes.</li> </ul>

<b>Automatización de Data Pipelines.</b>	
Competencias.	Actividades de aprendizaje.
<p><b>Específica(s):</b> Utilizar un servicio web para automatizar el movimiento y la transformación de los datos. Definir flujos de trabajo controlados por datos.</p>	<ul style="list-style-type: none"> <li>• Investigar los temas y conceptos claves sobre la construcción de data pipelines.</li> <li>• Realizar ejercicios aplicando en los data pipelines y data qualities.</li> </ul>



Número de registro: BRN-072  
Fecha de inicio: 2017-04-10  
Término de la certificación 2021-04-10

Av. Tecnológico S/N, Col. La Primavera, C.P. 47829 Ocotlán, Jalisco. Tel. 01 (392) 92 24630 y 92 53057, Ext. 120 e-mail: [aurora.nn@ocotlan.tecnm.mx](mailto:aurora.nn@ocotlan.tecnm.mx) [tecnm.mx](http://tecnm.mx) | [itocotlan.com](http://itocotlan.com)



**2022** Ricardo Flores  
Año de Magón  
PRECURSOR DE LA REVOLUCIÓN MEXICANA





<p>Genéricas:</p> <ul style="list-style-type: none"> <li>● Capacidad de análisis y síntesis.</li> <li>● Capacidad de aplicar los conocimientos en la práctica.</li> <li>● Capacidad de organizar y planificar.</li> <li>● Habilidad para buscar y analizar información proveniente de fuentes diversas.</li> <li>● Habilidades de investigación.</li> <li>● Capacidad de generar nuevas ideas.</li> <li>● Toma de decisiones.</li> <li>● Trabajo en equipo.</li> <li>● Liderazgo.</li> <li>● Habilidad para trabajar en forma. Autónoma.</li> <li>● Búsqueda del logro.</li> </ul>	<ul style="list-style-type: none"> <li>● Programar los elementos de un proyecto que el docente considere conveniente, de tal forma que el estudiante pueda hacer implementaciones de data pipelines.</li> </ul>
--	---

<b>Proyecto Final.</b>	
Competencias.	Actividades de aprendizaje.
<p>Específica(s): Implementar, justificar y desarrollar un proyecto final en común con otras asignaturas compatibles de acuerdo a la naturaleza del proyecto.</p> <p>Genéricas:</p> <ul style="list-style-type: none"> <li>● Capacidad de análisis y síntesis.</li> <li>● Capacidad de aplicar los conocimientos en la práctica.</li> <li>● Capacidad de organizar y planificar.</li> <li>● Habilidad para buscar y analizar información proveniente de fuentes diversas.</li> <li>● Habilidades de investigación.</li> <li>● Capacidad de generar nuevas ideas.</li> <li>● Toma de decisiones.</li> <li>● Trabajo en equipo.</li> <li>● Liderazgo.</li> <li>● Habilidad para trabajar en forma. Autónoma.</li> <li>● Búsqueda del logro.</li> </ul>	<ul style="list-style-type: none"> <li>● Implementar las tecnologías de ingeniería de los datos en un proyecto final en común con otras asignaturas compatibles.</li> <li>● Justificar la aplicación de las tecnologías utilizadas en el proyecto final.</li> <li>● Desarrollar la documentación que da soporte a la implementación de las tecnologías de ingeniería de los datos en el proyecto final.</li> <li>● Realizar una presentación de alto impacto en Blitz.</li> </ul>



Número de registro: 8911-072  
Fecha de inicio: 2017-04-10  
Término de la certificación 2021-04-10



## 8. Práctica(s).

- Implementar soluciones en los temas de unidad organizados en trabajo colaborativo.
- Visitar empresas que permitan la demostración del uso y aplicación de las tecnologías de ingeniería de datos en sus organizaciones.
- Verificar/decidir la plataforma en la nube a utilizar que que soporte los servicios y permita la implementación de las estrategias de ingeniería de datos especificadas en cada tema de unidad.
- Instalar las herramientas emergentes de desarrollo necesarias para el desarrollo de aplicaciones.
- Realizar aplicaciones a través de los lenguajes de scripting del lado cliente y servidor que permita poner en práctica el uso de tecnologías de ingeniería de datos.
- Realizar aplicaciones que pongan en práctica los temas de cada unidad.
- Programar las etapas de un proyecto de desarrollo y efectuar el control de versiones de diferentes instancias de código.
- Elaborar su portafolio digital y socializar en al menos una plataforma.

Nota: La definición específica de cada una de las prácticas, quedará a cargo del docente que imparta la materia y en función de los contenidos alcanzados.

## 9. Proyecto de asignatura.

El objetivo del proyecto que planteé el docente que imparta esta asignatura, es demostrar el desarrollo y alcance de la(s) competencia(s) de la asignatura, considerando las siguientes fases:

- **Fundamentación:** marco referencial (teórico, conceptual, contextual, legal) en el cual se fundamenta el proyecto de acuerdo con un diagnóstico realizado, mismo que permite a los estudiantes lograr la comprensión de la realidad o situación objeto de estudio para definir un proceso de intervención o hacer el diseño de un modelo.
- **Planeación:** con base en el diagnóstico en esta fase se realiza el diseño del proyecto por parte de los estudiantes con asesoría del docente; implica planificar un proceso: de intervención empresarial, social o comunitario, el diseño de un modelo, entre otros, según el tipo de proyecto, las actividades a realizar los recursos requeridos y el cronograma de trabajo.



- **Ejecución:** consiste en el desarrollo de la planeación del proyecto realizada por parte de los estudiantes con asesoría del docente, es decir en la intervención (social, empresarial), o construcción del modelo propuesto según el tipo de proyecto, es la fase de mayor duración que implica el desempeño de las competencias genéricas y específicas a desarrollar.
- **Evaluación:** es la fase final que aplica un juicio de valor en el contexto laboral-profesional, social e investigativo, ésta se debe realizar a través del reconocimiento de logros y aspectos a mejorar se estará promoviendo el concepto de “evaluación para la mejora continua”, la metacognición, el desarrollo del pensamiento crítico y reflexivo en los estudiantes.

## 10. Evaluación por competencias.

La evaluación debe ser permanente y continua. Se sugiere realizar una evaluación diagnóstica, formativa y sumativa. Así como aplicar autoevaluación, coevaluación y heteroevaluación.

Es pertinente la generación de un portafolio de evidencias, de preferencia en formato digital.

Algunas actividades de aprendizaje que se pueden utilizar:

- Organizadores gráficos como: Mapa conceptual, Mapa mental, entre otros.
- Guía de problemas de programación.
- Examen (teóricos y prácticos).
- Reportes de prácticas.
- Resúmenes.
- Cuadro sinóptico.
- Preguntas guiadas.
- Plenaria.
- Cuadro comparativo.
- Proyectos.
- Vídeos.
- Podcast.
- Screenshow.
- Desarrollo de un proyecto.
- Libros electrónicos.

Se pueden utilizar algunos instrumentos de evaluación como:

- Rúbrica.
- Lista de cotejo.
- Matriz de valoración.
- Guía de observación.
- Exámenes.



Número de registro: BRN-072  
Fecha de inicio: 2017-04-10  
Término de la certificación 2021-04-10



Av. Tecnológico S/N, Col. La Primavera, C.P. 47829 Ocotlán, Jalisco. Tel. 01 (392) 92 24630 y 92 53057, Ext. 120 e-mail: [aurora.nn@ocotlan.tecnm.mx](mailto:aurora.nn@ocotlan.tecnm.mx) [tecnm.mx](http://tecnm.mx) | [itocotlan.com](http://itocotlan.com)



**2022** Ricardo  
Flores  
Año de  
Magón  
PRECURSOR DE LA REVOLUCIÓN MEXICANA



## 11. Fuentes de información.

- [1]A. Gorelik, *The enterprise Big Data Lake*, 1st ed. O'Reilly Media, Inc., 2019.
- [2]R. Caballero and E. Martín Martín, *Las bases de Big Data*, 1st ed. Madrid: Los Libros de La Catarata, 2020.
- [3]M. Kleppmann, *Designing data-intensive applications*, 1st ed. Sebastopol, CA: O'Reilly, 2021.
- [4]R. Kitchin, *The data revolution*, 2nd ed. SAGE Publications Ltd; Second edition (May 31, 2022) Language : English, 2022.
- [5]A. Laurent, D. Laurent and C. Madera, *Data lakes*, 1st ed. Wiley-ISTE, 2020.
- [6]P. Crickard, *Data Engineering with Python*. Birmingham: Packt Publishing, Limited, 2020.
- [7]S. Akhtar, *Big Data Architect's Handbook*. Birmingham: Packt Publishing Ltd, 2018.
- [8]K. Gordon, *Principles of data management*. BCS, The Chartered Institute for IT, 2022.
- [9]V. Mayer-Schönberger, K. Cukier and A. Iriarte, *Big Data. La Revolución de los Datos Masivos*. Madrid: Turner, 2013.
- [10]"Big Data Specialization", *Coursera*, 2022. [Online]. Available: <https://www.coursera.org/specializations/big-data>. [Accessed: 29- Sep- 2022].
- [11]"IBM Data Engineering Professional Certificate", *Coursera*, 2022. [Online]. Available: <https://www.coursera.org/professional-certificates/ibm-data-engineer>. [Accessed: 29- Sep- 2022].
- [12]"Google Data Analytics Professional Certificate", *Coursera*, 2022. [Online]. Available: <https://www.coursera.org/professional-certificates/google-data-analytics>. [Accessed: 29- Sep- 2022].

