

MÁSTER EN

Ingeniería de datos

ONLINE

En EOI apostamos por un liderazgo consciente: sostenible, digital, diverso, emprendedor y transformador. Porque en EOI formamos en conocimiento, pero también en valores: perseguimos el bien común, el progreso y el equilibrio entre lo económico y humano. Porque en EOI **formamos talento para un futuro sostenible.**



www.eoi.es



Introducción

Todas las empresas están viviendo un ciclo de digitalización jamás visto hasta la fecha. Y eso implica la generación de una enorme cantidad de datos a los que hay que sacar valor. Pero desde los sistemas, webs y aplicaciones empresariales, hasta el reporting y la generación de algoritmos de inteligencia de artificial, hay un largo y duro camino, que marca la diferencia entre tomar buenas o malas decisiones, o que los algoritmos saquen conclusiones adecuadas que poder extrapolar las conductas de los clientes.

Dicho camino es el de la transformación y la adecuación del dato para los diferentes casos de uso. Y dicha adecuación necesita de muchas tecnologías, muy dependientes del caso de negocio a lograr. Todas estas tareas recaen bajo la responsabilidad del Data Engineer.

El objetivo de este máster es conocer todas las herramientas más habituales, tener un nivel alto en los lenguajes de programación necesarios para hacer ingeniería del dato, y ser capaces de implantar todos estos desarrollos en las arquitecturas adecuadas con los ciclos de implantación de software profesional que utilizan las grandes empresas, garantizando calidad, robustez y trazabilidad.

 **INICIO**
Marzo

 **DURACIÓN**
650 h

 **MODALIDAD**
Online + 1 semana
presencial en Madrid

 **SEDE**
EOI Madrid

 **PRECIO**
12.000 €



Fundación Estatal
PARA LA FORMACIÓN EN EL EMPLEO

Programa Bonificable
a través de FUNDAE.

Dirigido a

- Titulados superiores en Telecomunicaciones, Informática, Física, Matemáticas, ...
Para otras titulaciones se valora conocimiento y experiencia previos en programación de alto nivel, estadística, lenguaje SQL, ...
- **Ingenieros de datos con experiencia**, que vengan de entornos relacionales tipo SQL, y que necesitan adaptar sus conocimientos a las nuevas tecnologías.
Ecosistemas tecnológicos y lenguajes.
- **Ingenieros con poca experiencia laboral**, que necesitan especializarse en el mundo de los datos, o tener conocimientos para ser más empleables.



EOI tiene como propósito formar talento y, con el fin de facilitar el acceso a su oferta formativa al mayor número de personas posible, ha diseñado **el programa #eoiteayuda que contempla un amplio abanico de descuentos y bonificaciones que pueden llegar a cubrir el 40% del importe total del curso.**

Consulta las bases de esta convocatoria en www.eoi.es/programas/becas-y-descuentos

Tu futuro

El/la Ingeniero/a tiene un perfil muy técnico que se ocupa de administrar, gestionar, limpiar, procesar, organizar y enriquecer los datos, así como de implementar la infraestructura necesaria para almacenarlos de forma adecuada y así que estos se puedan utilizar de forma más eficiente por los analistas y científicos de datos para la toma de decisiones.

**Este programa
contribuye a la
Agenda 2030
en los siguientes ODS:**



Contenidos

Bloque 1. Arquitectura

INTRODUCCIÓN A LOS SISTEMAS INFORMACIONALES

- Evolución desde los sistemas operacionales a los sistemas informacionales.
- Tipos de sistemas informacionales.
- Perfiles profesionales.
- Dataflow, desde la creación hasta el caso de negocio.
- Data driven company, como sacar valor al dato.

VIRTUALIZACIÓN

- Orígenes de la virtualización. Tipos y ventajas.
- Máquinas virtuales.
- Contenedores, kubernetes y soluciones comerciales.
- Virtualización orientada al cómputo paralelo. Ejemplos comerciales.
- Virtualización orientada a los servicios.

ARQUITECTURA Y CLOUD

- Piezas tecnológicas que soportan el Dataflow.

- Diseños generales: Batch, lambda, kappa, etc.
- Alternativas OnPrem, opensource, comerciales.
- Cloudstack: AWS, GCS, Azure.

PROCESAMIENTO DISTRIBUIDO

- Conceptos de computación distribuida, cuellos de botella, hash.
- Conceptos básicos hadoop y distribuciones.
- HDFS, MapReduce, yarn.
- Futuro de hadoop, arquitectura cloudera y roadmap.

ALMACENAMIENTO Y COMPUTO ESCALABLE

- Tipos de bases de datos: mapa según tecnología, tipo de almacenamiento y acceso.
- BBDD SQL: Árbol, hash... Soluciones comerciales y open.
- BBDD NoSQL: HBase, Mongo, SolR, Neo4J, etc. Soluciones cloud native.

Bloque 2. Ingeniería del Dato

MODELO ENTIDAD RELACIÓN Y SQL BÁSICO

- Modelo relacional.
- Modelo Físico y lógico. Paso a tablas.
- SQL básico: SELECT, joins, agregaciones,

- subconsultas y operadores lógicos.
- Creación de tablas básico.

SQL AVANZADO

- Lenguaje de definición de datos.
- Lenguaje de manipulación avanzado: funciones ventana, índices, estadísticas, hints, macros y scripting.
- Lenguaje de control de datos: permisos y administración de objetos.

PYTHON Y PRINCIPIOS DE PROGRAMACIÓN

- Estructuras básicas de programación.
- Framework de desarrollo.
- Principios de Python.
- Tipos y estructuras de datos.
- Bloques y estructuras de control.
- Funciones y programación estructurada.
- E/S y gestión de archivos.
- Módulo y paquetes.

PYSPARK

- SparkShell.
- RDDs.
- Dataframes.
- Pandas, Koala.
- Spark SQL.
- Plan de ejecución y UI.

Contenidos

KAFKA Y SPARK STREAMING

- Definición de bus: Kafka / Confluent.
- Mensajes, Topics y el Log de mensajes.
- Productores, consumidores y brokers.
- Conectores.
- Streams.
- Particiones.
- Kafka Streams.
- Windowed Stream Processing.

PROCESOS PARA EXPLOTACION DE DATOS

- Estrella y copo de nieve.
- Cubos y OLAP.
- Desnormalización.
- Cálculo de KPIs complejos.

DESARROLLO PROCESOS ETL

- Extracción: replica, batch, microbatch, streaming, webscraping.
- Transformación y carga: procesos de carga way.
- Transformaciones básicas (SQL y Python).
- Monitorización, alarmado, control de flujo, calidad.

SPARK ORIENTADO A IA

- Frameworks: Notebooks.
- Muestreo, anonimización y filtrado.
- Estructuras tipo customer view.
- Normalización y discretización de variables.
- Funciones analíticas: MLlib & Others.
- Puesta en producción de un modelo analítico tipo batch.

CAPTURA DE LOGS Y MONITORIZACION

- Arquitectura ELK.
- Logstash.
- ElasticSearch.
- Kibana.
- Monitorización, alarmado y acceso al dato ElasticSearch.

APIFICACIÓN

- Tipos y securización: Rest, SOAP, apigateway, etc.
- API y virtualización.
- Desarrollo de API para DataService.
- Desarrollo de API para modelo analítico.
- Desarrollo de API en bases de datos de búsqueda.

Bloque 3. Gobierno E Implantación

CICD / MLOps

- Conceptos básicos de entornos de desarrollo (desarrollo, integración, producción, repositorio, etc).
- Ciclo CICD para desarrollos.
- Concepto de MLOps. Diferencias con CICD.
- Implantación Ciclo MLOps.
- Jenkins: recetas.
- Planificación de procesos.
- Ejemplo aplicación e2e.

GESTION Y GOBIERNO DEL DATO

- Data governance y data management.
- Metadata.
- Calidad, linaje, auditoría.



SESIONES ONLINE

La metodología online de EOI garantiza la misma calidad docente que el formato presencial, con un enfoque flexible e innovador que rompe con las barreras geográficas y la incompatibilidad de horarios. Fomentamos la interacción constante entre alumnos y profesores. El desarrollo del trabajo se distribuye según una rigurosa planificación y evaluación.

SEMANA PRESENCIAL EN ESPAÑA

Durante una estancia de una semana en Madrid, teoría y práctica se dan la mano mediante el estudio, análisis y desarrollo de multitud de supuestos y casos prácticos que permiten al participante consolidar sus conocimientos, orientarlos a la acción y vivir experiencias similares a las que experimenta en su desempeño profesional. Esta etapa formativa se compone de diversas actividades tales como:

- Ponencias
 - Conferencias y mesas redondas
 - Talleres
 - Trabajo en proyecto
-

PROYECTO FIN DE MÁSTER

Los alumnos deberán realizar un Proyecto Fin de Master al concluir el curso, que será presentado oralmente ante un tribunal. El proyecto consiste en la realización de un trabajo de aplicación práctico en grupo.

Los tutores del proyecto son profesores especialistas en los temas a desarrollar en cada proyecto.



Claustro

En EOI, los profesores constituyen una pieza fundamental del éxito de los cursos. El claustro está formado por directivos de empresa de reconocido prestigio profesional y comprobada trayectoria en la materia impartida, que transmiten no sólo los conocimientos teóricos, sino también su propia experiencia en la aplicación práctica de los mismos.

Los profesores te acompañarán durante todo el proceso de aprendizaje y estarán a tu disposición para consultas sobre cualquier tema de las materias que imparten.

Alberto Turégano

Asesor de compañías en proceso de transformación digital
LÍDERES Y DIGITALES

Alberto Oikawa

Director Técnico del Máster en Big Data & Business Analytics
CTO HISPATEC ANALYTICS

Jordi Conesa

Profesor
UNIVERSITAT OBERTA DE CATALUNYA

Ignacio Charfole

Head of Big Data Architecture and Delivery

Manuel Gómez

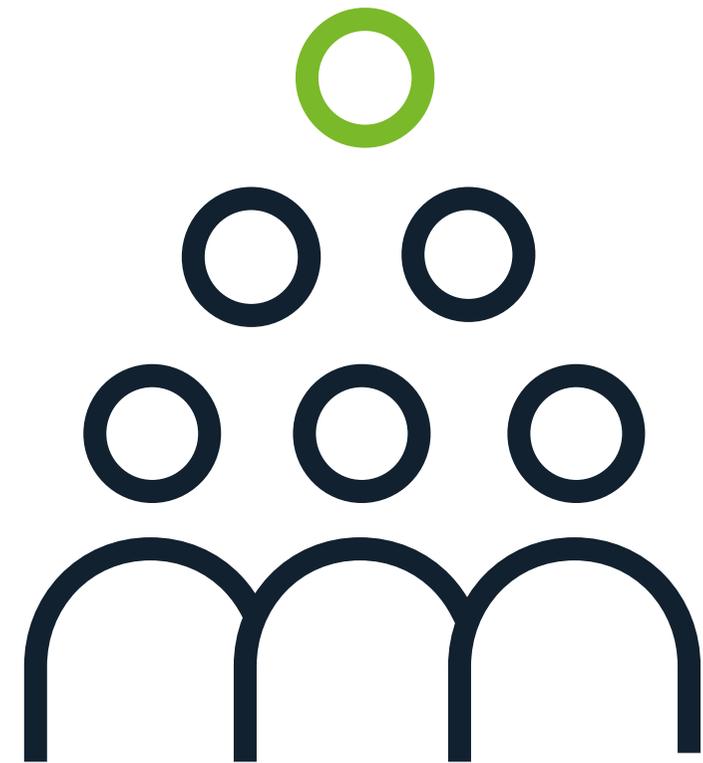
Arquitecto Big Data
TELEFÓNICA

Carlos Sacristán

Principal Data Architect
BIP SPAIN

Mario Renau

Senior Staff Data Engineer
ALTSTOM



Jose Novo

Data scientist & data analytics
O2 ESPAÑA

Rafael Zambrano

Head of Applied Science
OPENBANK

Daniel Burrueco

Business Intelligence & Machine Learning
CONSULTANT



Formamos **talento** para un futuro **Sostenible**



GOBIERNO
DE ESPAÑA

MINISTERIO
DE INDUSTRIA
Y TURISMO



Escuela de
organización
industrial

Madrid

informacion@eoi.es
+34 91 349 56 00
(Madrid)

Sevilla

infoandalucia@eoi.es
+34 95 446 33 77
(Sevilla)

Elche

eoimediterraneo@eoi.es
+34 96 665 81 55
(Elche - Alicante)

www.eoi.es

