

## Curso virtual

### Big Data

<b>INSTRUCTOR:</b>	Jorge Morato
<b>DURACIÓN:</b>	12 días (30 horas de trabajo)
<b>PERÍODO DE INSCRIPCIÓN:</b>	del 24 de junio al 25 de noviembre de 2017
<b>FECHAS DEL CURSO:</b>	7 al 16 de diciembre

#### INTRODUCCIÓN

El objetivo central del curso es ofrecer al alumno una aproximación global y práctica de big data, proporcionándoles los conocimientos necesarios para su comprensión así como las herramientas y estrategias que le permitan entender y rentabilizar el uso de los big data

Así mismo el curso busca proporcionar los fundamentos de la analítica de datos, analítica predictiva, los datos abiertos (open data) y la ciencia de datos soportes del científico de datos, la profesión “más sexy del siglo XXI, según la revista Harvard Business Review.

De igual forma el curso proporcionará los conocimientos básicos para entender las técnicas y herramientas clave en la manipulación de big data, tales como: Bases de datos analíticas (NoSQL y en memoria), marco de trabajo Hadoop, etc. y las aplicaciones de mayor impacto del Big Data en el periodismo de datos, ciudades inteligentes, internet de las cosas, salud, administración, etc.

#### OBJETIVOS DEL CURSO

Ofrecer al alumno una aproximación global y práctica de big data.

Comprender los fundamentos de la analítica de datos, analítica predictiva, los datos abiertos (open data) y la ciencia de datos soportes del científico de datos, la profesión “más sexy del siglo XXI, según la revista Harvard Business Review.

Entender las técnicas y herramientas clave en la manipulación de big data, tales como: Bases de datos analíticas (NoSQL y en memoria), marco de trabajo Hadoop, etc. y las aplicaciones de mayor impacto del Big Data en el periodismo de datos, ciudades inteligentes, internet de las cosas, salud, administración, etc.

## **METODOLOGÍA**

Los contenidos de esta asignatura estarán disponibles en la plataforma de teleformación virtual, que por su sencillez le permite al alumnado un aprendizaje flexible, acorde a sus necesidades.

El curso on-line se compone de varios módulos. El acceso a cada uno de estos módulos se efectuará de manera escalonada. El profesor establecerá en cada módulo un apartado teórico que irá acompañado de sus correspondientes actividades de aprendizaje como prácticas, test de evaluación y tutorías virtuales (foros, chat, correo electrónico).

La forma de comunicación profesor / alumno se efectuará en las dos modalidades: síncrona (chat) y asíncrona (foros de debate).

## **PERFIL DE LOS PARTICIPANTES**

A profesionales de áreas diversas como ciencias, ingeniería, documentación, comunicación, ciencias sociales, marketing, empresariales... estudiantes de master y doctorado así como a estudiantes universitarios (recomendable últimos años de grado) de iguales disciplinas que desean conocer los principios fundamentales de big data y de cómo se puede obtener rentabilidad y eficiencia a la manipulación de los grandes volúmenes de datos junto el conocimiento de las técnicas y herramientas utilizadas en organizaciones y empresas para utilizar con eficacia el poder de los grandes volúmenes de datos.

Así mismo el curso va dirigido a directivos de TIC, comerciales y de marketing, financieros y gerentes/analistas de consultoría que desea conocer una visión innovadora del uso de los datos y para el conocimiento adecuado para la toma de decisiones

## **CONTENIDO**

### **Modulo 1. COMPRENDIENDO BIG DATA**

Tema 1: Introducción: Marco social y tecnológico

1.1 La sociedad de las TIC

1.2 Papel actual de la información y los datos

1.3 Paradigmas de almacenamiento y proceso de datos. Evolución histórica

1.4 Caracterización del concepto Big Data

Tema 2: Fundamentación Teórica

2.1 Ciclo de Vida de la Información

2.2 Fuentes de la Información

## CURSOS VIRTUALES 2016

- 2.3 Estructuración y Saneamiento de los datos: la coherencia
- 2.4 Integración de fuentes
- 2.5 Arquitecturas de SI: centralizado vs. Distribuido: tolerancia a fallos
- 2.6 Resumen: el teorema CAP

### Tema 3: Aproximaciones a sistemas Big Data

- 3.1 BD Transaccionales (operacionales)
- 3.2 BD Analíticas
- 3.3 SQL, No-SQL, y No-Rel
- 3.4 Tipos de sistemas No-Rel y su conexión con BigData
- 3.5 Idoneidad Tecnológica

## **Modulo 2. APLICANDO BIG DATA**

### Tema 4: Implementando Big Data

- 4.1 Operatividad de BigData
- 4.2 El paradigma Map-Reduce
- 4.3 Implantación en organizaciones
- 4.4 Aspectos legales y éticos de Big Data. La Privacidad y la Seguridad de la Información

### Tema 5: Tecnologías Big Data

- 5.1 Hadoop
- 5.2 Mongo DB

### Tema 6: El Análisis de Datos

- 6.1 La analítica de datos
- 6.2 Análisis prospectivo
- 6.3 Análisis predictivo
- 6.4 Minería de Datos
- 6.5 Ciencia de Datos
- 6.6 Herramientas para el Análisis de datos (R, Pearl, ...)

## **Módulo 3. EXPLOTANDO BIGDATA**

### Tema 7: Big Data Ayer, Hoy y Mañana

- 7.1 Retrospectiva (ejemplos exitosos de aplicación de la tecnología)
- 7.2 Actuales Áreas de Negocio (Difusión de la Tecnología)
- 7.3 Perfiles Profesionales

#### PERFIL DEL INSTRUCTOR



**Jorge MORATO LARA**

España

#### **Español:**

Licenciado en Ciencias por la Universidad de Alcalá de Henares y Doctor en Documentación por la Universidad Carlos III. En el año 1990 comencé a trabajar en el CSIC y el INE. A partir del año 1994, con los cursos de doctorado, empiezo a investigar en la Universidad Carlos III. Desde el 97 tengo varias becas de investigación en el departamento de Informática de la Universidad Carlos III, que se convierten en el año 2000 en un contrato de profesor, desde el 2004 con el carácter permanente. Formo parte del departamento de informática desde hace 17 años, he impartido docencia también en otras titulaciones, en temas relacionados con la organización y recuperación de información, ingeniería del software y de la información, sobre todo en entornos web. Tengo reconocidos tres quinquenios de docencia así como reconocimientos (competitivos) por buenos resultados en docencia. He sido director de cuatro tesis, una de ellas premiada a la mejor tesis del año en el tema y otra con mención europea, y varias decenas de trabajos fin de grado y de máster. Poseo el certificado de aptitud pedagógica por la Universidad Complutense.

En 1999 presenté mi tesis en indicadores de calidad de textos médicos en lenguaje natural sobre extracción de relaciones semánticas. He participado en 4 proyectos nacionales y dos europeos. Así como en numerosos proyectos internacionales. He colaborado a su vez en números proyectos art. 83, siendo actualmente IP de cuatro. Tengo publicadas más de cien contribuciones. De entre las revistas catorce son JCR. He participado en cuatro proyectos de innovación docente y realizada estancias en las Universidades de Sao Paulo, Edimburgo, Ámsterdam, CNR de Italia, CNRS de Francia, Tesalónica y Miño. Para estas estancias tuve que participar siempre en procesos competitivos. He sido revisor de cuatro revistas JCR y miembro de organización de congresos nacionales e internacionales, comités nacionales de evaluación y

## CURSOS VIRTUALES 2016

órganos de gobierno universitario. Actualmente tengo dos sexenios de investigación reconocidos. Estoy acreditado por la ANECA para profesor titular. Desde el 2009 he recibido 9 reconocimientos docentes y he recibido 4 complementos por calidad de investigación o docencia convocados bianualmente por la Universidad.

Soy miembro de grupos como Ontolog, Instituto Universitario Julio Caro Baroja, CNPQ (Univ. Sao Paulo), Reusing Group y Bitrum. Además he firmado y sido coordinador Erasmus con las Universidades de Tempere, Reikiavik, Piura y Ámsterdam.