

## PLANTILLA GUÍAS DOCENTES Másteres Oficiales

### 1. DATOS GENERALES

- Breve descripción de la asignatura:
  - Esta materia tiene como objetivo proporcionar al estudiante un contacto directo con las tecnologías, principalmente en entorno de computación Cloud, que permiten el despliegue de las herramientas de análisis de datos y el desarrollo e implementación de nuevas soluciones.
- Título asignatura:
  - Herramientas en la nube para la Ciencia de Datos.
- Código asignatura:
  - M1973
- Curso académico:
  - 2019-2020
- Planes donde se imparte:
  - Máster universitario en ciencia de datos / Master in Sata Science.
- Créditos ECTS:
  - 4
- Carácter de la asignatura:
  - Optativa.
- Duración:
  - Cuatrimestral.
- Idioma:
  - Castellano e Inglés.

### • CONTENIDOS

- Principios básicos de computación Cloud
  - Fundamentos del modelo de computación Cloud.
  - Modelos de servicio: Infraestructura como Servicio (IaaS), Plataforma como servicio (PaaS), Software como servicio (SaaS).
  - Clouds híbridos. Acceso a recursos comerciales.
  - Composición de servicios. Herramientas básicas de orquestación.
  - Contenedores. Uso de Contenedores y Docker. Orquestación de contenedores.
- Sistemas de almacenamiento escalable de nueva generación.
  - Almacenamiento en entornos Cloud. Almacenamiento bloque. Almacenamiento de objeto.
  - Sistemas de almacenamiento para procesamiento paralelo.
  - Sistemas de analítica de datos avanzados para grandes volúmenes de datos.
  - Herramientas de almacenamiento.

### • COMPETENCIAS

- Generales:
  - Capacidad para integrarse eficazmente en un grupo de trabajo y trabajar en equipo, compartir la información disponible e integrar su actividad en la actividad del grupo colaborando de forma activa en la consecución de objetivos comunes.
  - Capacidad para redactar documentos científicos y técnicos, en particular artículos científicos.
  - Capacidad para planificar, diseñar y poner en marcha un proyecto avanzado.
  - Conocer las herramientas metodológicas necesarias para desarrollar proyectos avanzados.
  - Capacidad de actualización de los conocimientos expuestos en el ámbito de la comunidad científica.

- Transversales:
  - Capacidad para buscar, obtener, seleccionar, tratar, analizar y comunicar información utilizando diferentes fuentes
  - Capacidad para proyectar los conocimientos, habilidades y destrezas adquiridos para promover una sociedad basada en los valores de la libertad, la justicia, la igualdad y el pluralismo
  - Capacidad para afrontar tareas y situaciones críticas
  - Capacidades asociadas al trabajo en equipo: cooperación, liderazgo, saber escuchar
  - Motivación por la calidad
- Específicas:
  - Utilizar los principios de ingeniería para investigar, diseñar, desarrollar e implementar nuevos instrumentos y aplicaciones para la recogida de datos, su análisis y gestión.
  - Aplicar los principios de ingeniería a la investigación, diseño y desarrollo de un prototipo de aplicaciones de análisis de datos, o al desarrollo de estructuras, instrumentos, máquinas, experimentos, procesos, sistemas requeridos para ello.
  - Desarrollar y aplicar soluciones computacionales para problemas en un cierto dominio de aplicación, usando una amplia gama de plataformas de análisis de datos
  - Desarrollar e implementar herramientas de apoyo especializadas para el análisis de datos-.
  - Realizar prototipos de nuevas arquitecturas computacionales enfocadas a la analítica de datos.

#### 4. PLAN DE APRENDIZAJE

- Actividades formativas: descripción; horas; grado de presencialidad:
  - Dirigidas:
    - Participación y asistencia a lecciones magistrales en el aula
    - Participación y asistencia en seminarios dirigidos por un profesor
    - Realización de prácticas de computación y análisis de datos
  - Guiadas:
    - Desarrollo de proyectos guiados
    - Tutorías con un profesor que se desarrollarán tanto personalmente como por medio de recursos en red (por ejemplo, correo electrónico, gestor de contenidos en entorno web. e.g. Moodle)
  - Autónomas:
    - Elaboración de informes de laboratorio de datos
    - Realización y presentación escrita de trabajos
    - Estudio individual de contenidos de la asignatura
    - Estudio en grupo de contenidos de la asignatura
- Metodologías docentes:
  - En esta asignatura se comenzará por una exposición de los conceptos básicos, incluyendo ejemplos sencillos pero relevantes, que serán analizados individualmente y discutidos en común. Se revisarán los diferentes componentes de una solución, y los actores que participan en el desarrollo de la misma. Los estudiantes, organizados en grupos, realizarán una implementación práctica de algunos de los componentes descritos: análisis detallado, arquitectura e implementación final; por ejemplo el despliegue de una plataforma de orquestación de contenedores sobre un proveedor de infraestructura como servicio.
- Resultados de aprendizaje:
  - Identificar los recursos de computación adecuados, tanto locales como distribuidos, para abordar un problema de Data Science.
  - Definir los requerimientos técnicos de nuevas aplicaciones en analítica de datos a partir de un diseño de alto nivel.
  - Implementar algoritmos paralelos sencillos pero útiles en Data Science para su uso en un cluster o en un supercomputador.

- Saber diseñar una arquitectura Cloud Cloud en formato IaaS, PaaS o SaaS, para implementar una solución en Big Data
- Saber acceder a sistemas en entorno Cloud en formato IaaS, PaaS o SaaS, para implementar una solución en Big Data
- Utilizar una solución basada en contenedores (mediante la herramienta Docker).
- Cómo diseñar y gestionar un proyecto de computación Cloud.

## **5. SISTEMA DE EVALUACIÓN**

- Descripción del sistema de evaluación (ponderación mínima y ponderación máxima):
  - Examen escrito: 40%
  - Valoración de informes y trabajos: 60%
- Calendario de exámenes:
  - Al finalizar el curso.

## **6. PROFESORADO**

- Profesor responsable (firmante de actas):
  - Álvaro López García ([aloga@ifca.unican.es](mailto:aloga@ifca.unican.es)).
- Profesorado:
  - Pablo Orviz Fernández.
  - Iban Cabrillo Bartolomé.
  - Aida Palacio Hoz.
  - Fernando Aguilar.

## **7. SISTEMA DE EVALUACIÓN**

- Horario:
  - Tardes

## **8. BIBLIOGRAFÍA Y ENLACES RELACIONADOS**

- Bibliografía:
  - "Cloud Computing: Automating the Virtualized Data Center", 1st Ed. V. Josyula, M. Orr, G. Page. Cisco Press, 2012