# A14. Recuperación y extracción de información, grafos y redes sociales

**MÁSTER UNIVERSITARIO EN INVESTIGACIÓN EN INTELIGENCIA ARTIFICIAL**

***UNIVERSIDAD INTERNACIONAL MENÉNDEZ PELAYO***

Másteres Universitarios

**Este documento puede utilizarse como documentación de referencia de esta asignatura para la solicitud de reconocimiento de créditos en otros estudios. Para su plena validez debe estar sellado por la Secretaría de Estudiantes UIMP.**

# DATOS GENERALES

## Título asignatura

A14. Recuperación y extracción de información, grafos y redes sociales

## Código asignatura

102477

## Curso académico

2020-21

## Planes donde se imparte

MÁSTER UNIVERSITARIO EN INVESTIGACIÓN EN INTELIGENCIA ARTIFICIAL

## Créditos ECTS

4,5

## Carácter de la asignatura

OPTATIVA

## Duración

Anual

## Idioma

Castellano

# CONTENIDOS

## Contenidos

Debido a las cantidades ingentes de información disponible online, la web se ha convertido en un terreno habitual para el procesamiento de información (textual y multimedia) y la Ciencia de Datos. Lo mismo ocurre con los grafos, presentes en multitud de tareas/problemas y que presentan distintos retos en su análisis (redes sociales, interacciones biológicas, etc.).

En esta asignatura se persigue el procesamiento y fusión eficiente de este tipo de información no estructurada, abordando las técnicas de extracción y representación de información a partir de ella, para poder aplicar posteriormente procesos de Ciencia de Datos, procesamiento del lenguaje natural y/o visualización.

En el caso de los textos se abordará el uso de técnicas y métodos de procesamiento de información textual y minería web en distintas aplicaciones como análisis de textos, minería de textos, análisis de sentimientos, clasificación, etc.

En el caso de los grafos el objetivo principal consiste en el descubrimiento elementos o patrones representativos dentro de un grafo, con el fin de identificar conceptos que describen a las estructuras más importantes para una mejor interpretación de los datos. Se estudiarán procesos de minería de grafos tales como, emparejamiento de grafos, compresión de grafos, análisis de subgrafos, etc.

- Extracción y recuperación de información, clasificación y clustering de documentos.
- Textos.
- Imágenes.
- Emparejamiento de grafos, comprensión de grafos, análisis de subgrafos.
- Análisis de sentimientos, análisis de redes sociales.

**Unidades**

**1. Análisis de Redes Sociales**
Introducción a la Web 2.0 y al Análisis de Medios Sociales. Introducción a las Redes Complejas. Aplicaciones. Aspectos Básicos y Propiedades Estructurales de las Redes Complejas. Aspectos Básicos del Análisis de Redes Sociales. Medidas de Centralidad. Ejemplos de Aplicación.

**2. Detección de Comunidades**
Estructura de Comunidades. Justificación de la Necesidad de Detección. Métodos de Detección de Comunidades. Ejemplos de Aplicación.

**3. Difusión de Información en Redes Sociales**
Procesos Epidémicos y de Difusión. Modelos Clásicos de Propagación de Epidemias. Modelos de Difusión de Información en Redes. Ejemplos de Aplicación.

**4. Recuperación de Información**
Introducción. Procesado e indexación de documentos. Modelos de Recuperación de Información. Evaluación de la recuperación. Recuperación de información en la Web. Motores de búsqueda de código abierto. Técnicas avanzadas de RI.

**5. Minería de Texto y Análisis de Sentimientos**
Introducción. Minería de textos. Minería de datos en la Web (Minería de contenidos. Minería de opiniones y análisis de sentimientos. Minería de la estructura de la Web. Minería de uso de la Web).

**6. Clustering de Documentos**
Introducción al agrupamiento documental. Aplicaciones del agrupamiento. La base: "The cluster hypothesis". Componentes del proceso de agrupamiento.

# RESULTADOS DE APRENDIZAJE Y DE FORMACIÓN

## Generales

CG1 - Entender los conceptos, los métodos y las aplicaciones de la inteligencia artificial.

CG2 - Evaluar nuevas herramientas computacionales y de gestión del conocimiento en el ámbito de la Inteligencia Artificial.

## Específicas

CE3 - Seleccionar el mecanismo de representación del conocimiento y el método de razonamiento más adecuados al contexto donde serán utilizados y diseñar su aplicación para problemas en el ámbito de la Inteligencia Artificial.

CE4 - Conocer los principales modelos de razonamiento impreciso para valorar su adecuación a la resolución de problemas que surgen en el ámbito de la Inteligencia Artificial.

CE5 - Analizar las fuentes documentales propias del ámbito de la investigación en Inteligencia Artificial para poder determinar cuáles de ellas son relevantes en la resolución de problemas concretos.

# PLAN DE APRENDIZAJE

## Actividades formativas

A1 - **Sesiones presenciales virtuales (clases en vídeo)**: visionado inicial del material audiovisual que constituye las lecciones de la asignatura. Se asume 1.5 veces el tiempo real de vídeo, puesto que el estudiante deberá parar, repetir, etc. algunas secuencias (18 horas).

A2 - **Trabajos individuales**: realización de ejercicios, resolución de problemas, realización de prácticas y/o trabajos/proyectos individuales (50 horas).

A3 - **Trabajo autónomo**: estudio del material básico, lecturas complementarias y otros contenidos (32,5 horas).

A4 - **Foros y chats**: lanzamiento, lectura y contestación de cuestiones y temas para la discusión general (6 horas).

A5 - **Tutorías**: consultas y resolución de dudas, aclaraciones, etc. (6 horas).

Puede consultar en este enlace el **Cronograma de Carga de Trabajo**.

# SISTEMA DE EVALUACIÓN

## Descripción del sistema de evaluación

E1 - **Valoración de los cuestionarios de evaluación**: los estudiantes realizarán un cuestionario de evaluación por cada una de las seis unidades didácticas. Dichos cuestionarios serán objeto de puntuación en la nota final. El peso en la nota final de este apartado será del 30% sobre el total (5% para cada cuestionario).

E2 - **Valoración de la participación en foros y chats**: se valorará el nivel de participación/debate de los estudiantes que contará para la nota final. El peso en la nota final de este apartado será del 10% sobre el total.

E3 - **Valoración de los trabajos individuales**: los alumnos deberán presentar un trabajo individual sobre cada bloque temática de la asignatura (Bloque I: Redes Sociales y Bloque II: Recuperación de Información y Minería de Texto) basados en casos de estudio propuesto por los profesores, en los que se demuestren las competencias adquiridas durante la asignatura. Su valoración corresponde al 60% del total de la asignatura (30% para cada trabajo).

## Calendario de exámenes

Para la **convocatoria ordinaria**, habrá 3 fechas de entrega de trabajos final de curso. Los alumnos podrán entregar sus trabajos en cualquier momento, pero sólo en estas fechas se recogerán y evaluarán los que se hayan entregado. Las fechas serán:

- 20/12/19
- 15/03/20
- 31/05/20

Habrá una **convocatoria extraordinaria** en todas las asignaturas. Para su evaluación, la fecha límite para la entrega de trabajos será:

- 10/07/20

Para los **Trabajos Fin de Máster** habrá dos convocatorias:

- Convocatoria ordinaria: Entrega de TFM hasta el 01/07/20 y defensa el 15/07/20
- Convocatoria extraordinaria: Entrega de TFM hasta el 01/09/20 y defensa el 15/09/20

Las actas de la convocatoria ordinaria se cerrarán en julio de 2020 y las de la convocatoria extraordinaria en septiembre de 2020.

# PROFESORADO

## Profesor responsable

**Cordón García, Óscar**

*DOCTOR EN INFORMÁTICA.*
*CATEDRÁTICO DE UNIVERSIDAD.*
*UNIVERSIDAD DE GRANADA - DPTO. CIENCIAS DE LA COMPUTACIÓN E INTELIGENCIA ARTIFICIAL.*

## Profesorado

**Fernández Luna, Juan Manuel**

*Licenciado en Informática, Doctor en Informática.*
*Catedrático de Universidad del área de Ciencia de la Computación e Inteligencia Artificial.*
*Universidad de Granada.*

**Chica Serrano, Manuel**

*Teaching Assistant*
*Investigador en Ramón y Cajal*
*Universidad de Granada*

# HORARIO

## Horario

Todas las asignaturas estarán en la plataforma a disposición de los estudiantes desde octubre hasta julio.

# BIBLIOGRAFÍA Y ENLACES RELACIONADOS

## Bibliografía

### Bibliografía fundamental

Zafarani, R., Ali Abbasi, M., Liu, H., *Social Media Mining. An Introduction*. Cambridge University Press 2014.

Wasserman, S., Faust, K. *Social Network Analysis. Methods and Applications*. Cambridge University Press 1994. ISBN 9780521387071.

Cacheda, F., Fernández-Luna, J.M. y Huete, J. *Recuperación de Información: Un enfoque práctico y multidisciplinar*. Ed. Rama 2011.

Bing Liu, *Sentiment Analysis: mining opinions, sentiments, and emotions*. Cambridge University Press, 2015.

### Bibliografía complementaria

Albert-Laszlo Barabasi: "Network Science". Interactive Book Project, http://barabasilab.com/networksciencebook

Mark Newman: "Networks: An introduction", Oxford, 2010. ISBN 0199206651.

Kumar, S., Morstatter, F., Liu, H. *Twitter Data Analytics*. Springer 2013.

McCandless, M., Hatcher, E., Gospodnetic, O. *Luce in Action* (2d Edition). Manning 2011.

### Enlaces recomendados

**Sobre Redes Complejas**:

Libro electrónico "Network Science Interactive Book Project": http://barabasilab.com/networksciencebook

Web de Mark Newman, University of Michigan: http://www-personal.umich.edu/~mejn/

Web de Réka Albert, Pennsylvania State University: http://www.phys.psu.edu/~ralbert/

**Sobre Análisis de Redes y Medios Sociales**:

Stanford Network Analysis Project: http://snap.stanford.edu/

Grupo "Social Mining", Max Planck Institute: http://ella.slis.indiana.edu/~katy/S637-S11/

**Sobre Recuperación de Información**:

Libro en línea sobre RI: http://npl.stanford.edu/IR-book

Web del grupo de IR de la ACM: http://www.acm.org/sigir