

Academic Year: (2022 / 2023)

Review date: 02-06-2022

Department assigned to the subject: Library and Information Sciences Department

Coordinating teacher: GARCIA ZORITA, JOSE CARLOS

Type: Electives ECTS Credits : 3.0

Year : 1 Semester : 0

REQUIREMENTS (SUBJECTS THAT ARE ASSUMED TO BE KNOWN)

- Database Design (340-17452).
- Statistical Data Analysis (340-17467).
- Data Visualisation (340-17468).
- Metric Studies of Information (340-17461).
- Data Science (340-17477).
- Information Visualisation (335-17283).

OBJECTIVES**BASIC COMPETENCES**

CB8 Learn how to design strategies for the analysis and exploitation of large volumes of data.

CB9 Make proposals for data management for different contexts and organisations.

GENERAL COMPETENCES

CG2 Learn to identify work and research lines related to the analysis, cleansing and exploitation of data.

CG6 Know the different business models related to Big data.

CG8 Learn to identify the potential value and use of data.

CG9 Learn how to adapt and develop methods and techniques according to the new knowledge and skills required for data management and analysis.

LEARNING RESULTS

- Knowledge of methods and techniques used for designing and assessing strategies of data management.

The student after passing the subject must:

- Apply techniques to elaborate studies and reports analysing data management in organisations.
- Know and understand concepts and terms related to big data.
- Design, plan and implement data cleansing and analysis processes.
- Learn to identify strategic data and high impact in the organisation.
- Know the different methods and techniques of data extraction, data cleansing, and linked data.
- Acquire a good command using tools for data extraction, data cleansing, and linked data.

The student, after passing the subject, must:

- Know the limitations and technical capabilities required in Big Data.
- Understand and understand the value of data from different sources.
- Knowing tools that allow designing, preparing, analysing and managing large volumes of structured or unstructured information.
- Use data analysis techniques to obtain valid conclusions for decision making.
- Normalise, relate and enrich the information provided by different datasets.
- Use different techniques and tools for data cleansing, data analysis and linking data.

DESCRIPTION OF CONTENTS: PROGRAMME

Module 1.- Introduction to Big data. Basic fundamentals. Concepts

1.1. Big data in the context of records management.

1.2. New trends in data management. Big data and Smart data

- 1.3. Big data in the strategies of public and private organizations
- 1.4. Trends, policy, issues and initiatives

Module 2.- Data analysis and filtering techniques.

- 2.1. Data quality audit
- 2.2. Structured and unstructured data testing tools. Introduction to Open Refine.
- 2.3. Unstructured data mining
- 2.4. Predictive models in data mining

Module 3.- Data processing and open linked data

- 3.1. Linked data, interoperability and the semantic web standards
- 3.2. Database connectivity using WebService
- 3.3. Data Extraction. Open source solutions
- 3.4. Project planning and data Processing. Case study

LEARNING ACTIVITIES AND METHODOLOGY

TRAINING ACTIVITIES

- AF1 Individual work for the study of theoretical and practical materials.
- AF2 practical cases.
- AF3 Theoretical-practical classes.
- AF4 Tutorship.
- AF5 Final work.
- AF6 Active participation in fora of the course.

TEACHING METHODS

- MD1 Oral presentations describing the key concepts of the subject.
- MD2 Critical reading of texts recommended.
- MD3 Resolution of practical exercises individually or in a group.
- MD4 Class discussions related to the content of the subject and/or practical cases.
- MD5 Preparation of individual and group work and reports.
- MD6 Reading of theoretical and practical teaching materials.

TUTORSHIP SCHEME

The schedules of the tutorship are available at Aula Global. In addition to these official tutorship, students may request and arrange additional tutorship with the teacher.

ASSESSMENT SYSTEM

EVALUATION SYSTEM

Tests and assignments	50%
Final exam (*)	50%

(*) The final exam will be held in face-to-face mode at the Carlos III University of Madrid, and must be passed to pass the subject.

EXTRAORDINARY CALL

In this subject, in the absence of specific university regulations for postgraduate studies, if the student has not followed the continuous assessment, he will be entitled to a test that will allow him to obtain 60% of the final maximum qualification in the exam Ordinary In the extraordinary, the exam will allow you to get up to 75% of the final maximum score.

% end-of-term-examination:	50
% of continuous assessment (assignments, laboratory, practicals...):	50

BASIC BIBLIOGRAPHY

- CABALLERO, Rafael y MARTIN, Enrique Las bases del Big Data, Madrid: CATARATA, 2017
- GÓMEZ GARCÍA, José Luis Introducción al big data, Barcelona: UOC, 2015

- JOYANES AGUILAR, Luis Big data: análisis de grandes volúmenes de datos en organizaciones, Barcelona: Marcombo, 2013
- LARA TORRALBO, Juan Alfonso Minería de datos, Madrid: CEF, 2014
- MARR, Bernard Big Data. La utilización del Big Data, el análisis y los parámetros SMART para tomar mejores decisiones y aumentar el rendimiento, Madrid: TEEL, 2016
- MARR, Bernard Big data en la práctica, Madrid: TEEL, 2017
- MAYER-SCHÖNBERGER, Viktor Big data: la revolución de los datos masivos, Madrid: Turner, 2013
- NETTLETON, David F Data mining: fundamentos y metodologías, Barcelona: UOC, 2007
- SCHMARZO, Bill Big data: el poder de los datos, Madrid: Anaya Multimedia, 2014
- SIEGEL, Eric Analítica predictiva: predecir el futuro utilizando Big Data, Madrid: Anaya Multimedia, 2013