ANX-PR/CL/001-01

LEARNING GUIDE

SUBJECT

103000686 - Large scale data management

DEGREE PROGRAMME

10AP - Eit Digital Master's Programme In Data Science

ACADEMIC YEAR & SEMESTER

2018/19 - Semester 1

Index

Learning guide

1. Description

1.1. Subject details

Name of the subject	103000686 - Large scale data management
No of credits	4.5 ECTS
Type	Compulsory
Academic year ot the programme	First year
Semester of tuition	Semester 1
Tuition period	September-January
Tuition languages	English
Degree programme	10AP - Eit digital master's programme in data science
Centre	10 - Escuela Tecnica Superior de Ingenieros Informaticos
Academic year	2018-19

2. Faculty

2.1. Faculty members with subject teaching role

Name and surname	Office/Room	Email	Tutoring hours *
Tonghong Li	2312	tonghong.li@upm.es	M - 14:00 - 16:00 W - 14:00 - 16:00 Th - 14:00 - 16:00
Marta Patiño Martinez (Subject coordinator)	2313	marta.patino@upm.es	Tu - 12:00 - 14:00 Th - 10:00 - 12:00 Th - 14:00 - 15:00 Please, send an email in order to minimize waiting time

* The tutoring schedule is indicative and subject to possible changes. Please check tutoring times with the faculty member in charge.

2.2. Research assistants

Name and surname	Email	Faculty member in charge
Vianello ., Valerio	valerio.vianello@upm.es	Patiño Martinez, Marta

2.3. External faculty

Name and surname	Email	Institution
Ainhoa Azqueta	aazqueta@fi.upm.es	UPM

3. Prior knowledge recommended to take the subject

3.1. Recommended (passed) subjects

El plan de estudios Eit Digital Master's Programme In Data Science no tiene definidas asignaturas previas recomendadas para esta asignatura.

3.2. Other recommended learning outcomes

- Java programming, concurrent programming, databases

4. Skills and learning outcomes *

4.1. Skills to be learned

CE01 - Capacidad para la integración de tecnologías, aplicaciones, servicios y sistemas propios de la Ingeniería Informática, con carácter generalista, y en contextos más amplios y multidisciplinares.

CE09 - Capacidad para diseñar y evaluar sistemas operativos y servidores, y aplicaciones y sistemas basados en computación distribuida.

CE11 - Capacidad de diseñar y desarrollar sistemas, aplicaciones y servicios informáticos en sistemas empotrados y ubicuos.

CG02 - Capacidad de gestionar la información

4.2. Learning outcomes

RA14 - Diseñar y evaluar sistemas operativos y servidores

RA35 - Diseñar e implementar sistemas altamente paralelos y/o distribuidos

RA32 - Conocer y saber utilizar técnicas fundamentales de computación de altas prestaciones.

* The Learning Guides should reflect the Skills and Learning Outcomes in the same way as indicated in the Degree Verification Memory. For this reason, they have not been translated into English and appear in Spanish.

5. Brief description of the subject and syllabus

5.1. Brief description of the subject

This course presents architectures for scalable distributed systems and data management systems: map-reduce,bigtable, data streaming, persistent queues

5.2. Syllabus

1. Introduction

2. Data management technologies

3. Data Streaming

4. Big Table. Dynamo

5. Contenedores. AWS

6. Schedule

6.1. Subject schedule*

Week	Face-to-face classroom activities	Face-to-face laboratory activities	Other face-to-face activities	Assessment activities
1	**Introducción** Duration: 02:00 Lecture			
2	**Tema 1** Duration: 03:00 Lecture			
3	**Tema 1** Duration: 03:00 Lecture			
4	**Tema 1** Duration: 03:00 Lecture			
5	**Tema 2** Duration: 03:00 Lecture			
6	**Tema 2** Duration: 02:30 Lecture			
7	**Tema 3** Duration: 03:00 Lecture			**Practical Assignament** Group work Final examination Duration: 10:00 **Practical assignament** Problem-solving test Continuous assessment Duration: 10:00
8	**Tema 3** Duration: 03:00 Lecture			
9	**Tema 4** Duration: 03:00 Lecture			
10	**Tema 4** Duration: 03:00 Lecture			
11	**Tema 5** Duration: 03:00 Lecture			
12	**Tema 5** Duration: 03:00 Lecture			

13	**Tema 6** Duration: 03:00 Lecture			
14	**Tema 6** Duration: 03:00 Lecture			
15	**Tema 6** Duration: 03:00 Lecture			**Practical assignament** Group work Final examination Duration: 10:00 **Practical assignment** Problem-solving test Continuous assessment Duration: 10:00
16				
17				**Exam** Written test Continuous assessment and final examination Duration: 00:00

The independent study hours are training activities during which students should spend time on individual study or individual assignments.

Depending on the programme study plan, total values will be calculated according to the ECTS credit unit as 26/27 hours of student face-to-face contact and independent study time.

* The subject schedule is based on a previous theorical planning of the subject plan and might go to through experience some unexpected changes along throughout the academic year.

7. Activities and assessment criteria

7.1. Assessment activities

7.1.1. Continuous assessment

Week	Description	Modality	Type	Duration	Weight	Minimum grade	Evaluated skills
7	Practical assignament	Problem-solving test	Face-to-face	10:00	35%	5 / 10	CG02 CE09 CE11 CE01
15	Practical assignment	Problem-solving test	Face-to-face	10:00	35%	5 / 10	CE09 CE11 CE01 CG02
17	Exam	Written test	Face-to-face	00:00	30%	4 / 10	CG02 CE09 CE11 CE01

7.1.2. Final examination

Week	Description	Modality	Type	Duration	Weight	Minimum grade	Evaluated skills
7	Practical Assignament	Group work	Face-to-face	10:00	35%	5 / 10	CG02 CE09 CE11 CE01
15	Practical assignament	Group work	Face-to-face	10:00	35%	5 / 10	CG02 CE09 CE11 CE01
17	Exam	Written test	Face-to-face	00:00	30%	4 / 10	CG02 CE09 CE11 CE01

7.1.3. Referred (re-sit) examination

Description	Modality	Type	Duration	Weight	Minimum grade	Evaluated skills
assigment	Group work	Face-to-face	10:00	35%	5 / 10	
Assignament	Group work	Face-to-face	10:00	35%	5 / 10	
Exam	Written test	Face-to-face	03:00	30%	4 / 10	

7.2. Assessment criteria

Assignments 70%

Exam 30%

8. Teaching resources

8.1. Teaching resources for the subject

Name	Type	Notes
Bibliografía	Bibliography	NoSQL Distilled: A Brief Guide to the Emerging World of Polyglot Persistence. P. Sadalage, M. Fowler. 2012.
Libro	Bibliography	Big Data Now: Current Perspectives from O'Reilly Radar. O?Reilly. 2011
libro2	Bibliography	Graph Databases. I. Robinson, J. Webber, E. Eifrem. O?Really. 2013
Slides	Bibliography	Slides
Papers	Bibliography	List of papers to be provided