



POLITÉCNICA

INTERNATIONAL
CAMPUS OF
EXCELLENCE

COORDINATION PROCESS OF
LEARNING ACTIVITIES
PR/CL/001



E.T.S. de Ingenieros
Informáticos

ANX-PR/CL/001-01

LEARNING GUIDE

SUBJECT

103000387 - Cloud Computing And Big Data Ecosystems Design

DEGREE PROGRAMME

10AK - Master Universitario En Software Y Sistemas

ACADEMIC YEAR & SEMESTER

2023/24 - Semester 1

Index

Learning guide

| | |
|---|---|
| 1. Description..... | 1 |
| 2. Faculty..... | 1 |
| 3. Prior knowledge recommended to take the subject..... | 2 |
| 4. Skills and learning outcomes | 2 |
| 5. Brief description of the subject and syllabus..... | 4 |
| 6. Schedule..... | 6 |
| 7. Activities and assessment criteria..... | 8 |
| 8. Teaching resources..... | 9 |

1. Description

1.1. Subject details

| | |
|---------------------------------------|--|
| Name of the subject | 103000387 - Cloud Computing And Big Data Ecosystems Design |
| No of credits | 4 ECTS |
| Type | Optional |
| Academic year of the programme | First year |
| Semester of tuition | Semester 1 |
| Tuition period | September-January |
| Tuition languages | English |
| Degree programme | 10AK - Master Universitario en Software y Sistemas |
| Centre | 10 - Escuela Tecnica Superior De Ingenieros Informaticos |
| Academic year | 2023-24 |

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 Please, write an email to agree on a date |
| Marta Patiño Martínez (Subject coordinator) | 2313 | marta.patino@upm.es | Tu - 12:00 - 14:00 Th - 10:00 - 12:00 Th - 14:00 - 15:00 Please, write an email to agree on a |

| | | | |
|-----------------------|--|-----------------------|---|
| | | | date |
| Ainhoa Azqueta Alzuaz | | ainhoa.azqueta@upm.es | Sin horario. Please, write an email to agree on a date |

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

3. Prior knowledge recommended to take the subject

3.1. Recommended (passed) subjects

The subject - recommended (passed), are not defined.

3.2. Other recommended learning outcomes

- Java programming, concurrent programming, databases

4. Skills and learning outcomes *

4.1. Skills to be learned

CEM2 - Analizar y sintetizar soluciones a problemas que requieran aproximaciones novedosas para la definición de la infraestructura computacional que permita el procesamiento y el análisis de datos de diversa naturaleza

CEM4 - Analizar y evaluar los diferentes paradigmas y enfoques de ingeniería de construcción y gestión de sistemas basados en software.

CEM6 - Realizar trabajos de investigación en las principales líneas de investigación activas en el área de los paradigmas de la computación distribuida, sus aplicaciones prácticas y la gestión de la infraestructura necesaria

CG12 - Comprensión amplia de las técnicas y métodos aplicables en una especialización concreta, así como de sus límites

CG13 - Apreciación de los límites del conocimiento actual y de la aplicación práctica de la tecnología más reciente.

CG14 - Conocimiento y comprensión de la informática necesaria para la creación de modelos de información, y de los sistemas y procesos complejos

CG4 - Que los estudiantes posean las habilidades de aprendizaje que les permitan continuar estudiando de un modo que habrá de ser en gran medida autodirigido o autónomo.

CG7 - Especificación y realización de tareas informáticas complejas, poco definidas o no familiares

CG9 - Aplicación de los métodos de resolución de problemas más recientes o innovadores y que puedan implicar el uso de otras disciplinas

CG120 - Adquirir conocimientos científicos avanzados del campo de la informática que le permitan generar nuevas ideas dentro de una línea de investigación.

CG123 - Capacidad de leer y comprender publicaciones dentro de su ámbito de estudio/investigación, así como su catalogación y valor científico

4.2. Learning outcomes

RA51 - Ser capaz de estudiar y analizar un nuevo área de investigación en sistemas distribuidos

RA50 - Ser capaz de encontrar e identificar artículos seminales de un área de investigación en sistemas distribuidos

RA104 - RA49

RA60 - Ser capaz de entender artículos de investigación en el área de sistemas distribuidos

RA58 - Conocer principales protocolos de control de concurrencia y recuperación

RA57 - Conocer principios sistemas transaccionales

RA53 - Ser capaz de identificar los principales retos que aborda una línea de investigación emergente en sistemas distribuidos

RA52 - Ser capaz de relacionar una línea de investigación emergente en sistemas distribuidos con los fundamentos de los sistemas distribuidos

RA49 - Conocer principales líneas activas en sistemas distribuidos

RA59 - Conocer principales protocolos de replicación de datos y de procesos y criterios de corrección

RA109 - RA57

RA105 - RA50

RA106 - RA51

RA107 - RA52

RA110 - RA59

RA111 - RA60

RA108 - RA53

RA56 - Conocer el radiado fiable: tipos, propiedades y protocolos

RA54 - Conocer caracterización sincronía sistemas distribuidos e implicaciones en posibilidad/imposibilidad de resolución de protocolos de acuerdo y coordinación

RA55 - Conocer principales protocolos de coordinación y acuerdo

* 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: bigtable, data streaming, persistent queues

5.2. Syllabus

1. Introduction
2. Data management technologies
 - 2.1. NoSQL: key-value, graph databases, document oriented databases
 - 2.2. SQL y NewSQL: column oriented data stores
 - 2.3. Complex Event Processing/Data streaming
3. Data Streaming
4. Big Table. Dynamo
5. Diseño Ecosistemas Gestión de Datos Big Data y Cloud

6. Schedule

6.1. Subject schedule*

| Week | Classroom activities | Laboratory activities | Distant / On-line | Assessment activities |
|------|--|-----------------------|-------------------|-----------------------|
| 1 | Introduction Duration: 02:00 Lecture | | | |
| 2 | Tema 1 Duration: 02:00 Lecture | | | |
| 3 | Tema 1 Duration: 02:00 Lecture | | | |
| 4 | Tema 1 Duration: 02:00 Lecture | | | |
| 5 | Tema 2 Duration: 02:00 Lecture | | | |
| 6 | Tema 2 Duration: 02:00 Lecture | | | |
| 7 | Tema 3 Duration: 02:00 Lecture | | | |
| 8 | Tema 3 Duration: 02:00 Lecture | | | |
| 9 | Tema 4 Duration: 02:00 Lecture | | | |
| 10 | Tema 4 Duration: 02:00 Lecture | | | |
| 11 | Tema 5 Duration: 02:00 Lecture | | | |
| 12 | Tema 5 Duration: 02:00 Lecture | | | |
| 13 | Exercices Duration: 02:00 Problem-solving class | | | |
| 14 | Assingment presentation Duration: 02:00 Additional activities Assingment presentation Duration: 02:00 Additional activities Assingment presentation | | | |

| | | | | |
|----|--|--|--|--|
| | Duration: 02:00 Additional activities | | | |
| 15 | Assingment presentation Duration: 02:00 Additional activities Assingment presentation Duration: 02:00 Additional activities | | | |
| 16 | | | | Exam Written test Continuous assessment Presential Duration: 03:00 Exam Written test Final examination Presential Duration: 03:00 |
| 17 | | | | |

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 schedule is based on an a priori planning of the subject; it might be modified during the academic year, especially considering the COVID19 evolution.

7. Activities and assessment criteria

7.1. Assessment activities

7.1.1. Assessment

| Week | Description | Modality | Type | Duration | Weight | Minimum grade | Evaluated skills |
|------|-------------|--------------|--------------|----------|--------|---------------|---|
| 16 | Exam | Written test | Face-to-face | 03:00 | 100% | 5 / 10 | CEM2 CG4 CG7 CG9 CG12 CG14 CGI20 CGI23 CEM4 CEM6 |

7.1.2. Global examination

| Week | Description | Modality | Type | Duration | Weight | Minimum grade | Evaluated skills |
|------|-------------|--------------|--------------|----------|--------|---------------|---|
| 16 | Exam | Written test | Face-to-face | 03:00 | 100% | 5 / 10 | CEM2 CG4 CG7 CG9 CG12 CG14 CGI20 CGI23 CEM4 CEM6 |

7.1.3. Referred (re-sit) examination

| Description | Modality | Type | Duration | Weight | Minimum grade | Evaluated skills |
|-------------|--------------|--------------|----------|--------|---------------|---------------------------------------|
| Exam | Written test | Face-to-face | 03:00 | 100% | 5 / 10 | CEM2 CG4 CGI20 CGI23 CEM6 |

7.2. Assessment criteria

Regular period:

Exam 100%

Solo prueba final:

Exam 100%

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'Reilly. 2013 |

| | | |
|--------|--------------|-------------------------------|
| Slides | Bibliography | Slides |
| Papers | Bibliography | List of papers to be provided |