

Degree programme	Type	Course
Data Engineering	OP	4

Contact lecturer

Name : Daniel Franco Puentes

Email : daniel.franco@uab.cat

Teaching staff

Antonio Espinosa Morales

Andreu Perez Martinez

Alvaro Wong Gonzalez

Group languages

You can consult this information at the [end](#) of the document.

Prerequisites

This subject has no PRE-REQUIREMENTS. It is recommended to have completed the following subjects:

[104357 - Computation in Cloud Environments](#)

[104365 - Data Visualisation](#)

[104358 - Development of Big Data Applications](#)

[104362 - Neural Networks and Deep Learning \(2021-22\)](#)

Objectives

The objective of this course is to know the advanced methods and concepts of processing and management of massive data, both from the point of view of generation, adaptation, transmission and storage and, also, the processing and analysis to extract useful information. It is also a goal to use the right tools to carry out the work with massive data, both interactively and locally as well as batch and remote and deferred and real-time.

Learning outcomes

1. Students must be capable of collecting and interpreting relevant data (usually within their area of study) in order to make statements that reflect social, scientific or ethical relevant issues.
2. Work cooperatively in complex and uncertain environments and with limited resources in a multidisciplinary context, assuming and respecting the role of the different members of the group.
3. Prevent and solve problems, adapt to unforeseen situations and take decisions.
4. Students must develop the necessary learning skills to undertake further training with a high degree of autonomy.
5. Study the adaptations made to data analysis and consultation algorithms in order for these to maintain the privacy of the entry data, of the models learnt, or of the outputs of the models used in the field of business intelligence.

Contents

1. Advanced methods of big data processing and management
2. Reference Cloud architectures for big data management and application design: batch, streaming, decentralized.
3. Analysis of challenges and case studies with cloud-oriented and big data
4. MVP methodology for the design and development of Cloud solutions to manage big data
5. Tools for evaluating applications and big data management systems

Learning activities and methodology

Title	Hours	ECTS	Learning outcomes
Self-study	100	4	1, 2, 3, 4, 5
Analysis, design and development of prototypes	25	1	1, 2, 3, 4, 5
Lectures on technology, methodology and the case study	20	0.8	1, 5

→ Teaching methodology

The subject proposes a work with the Challenge-Based Learning (ABR) methodology that consists of a real challenge posed by an entity/institution with which it will have a relationship and will require teamwork to propose possible solutions, which are planned and developed in three phases: Bond/Commitment, Research/Prototyping and Implementation/Evaluation.

→ Training activities

Methodologies will be used to work based on challenges (ABR) that will use various activities such as: reading articles, conferences, document analysis, surveys, interviews, bibliographic searches, presentations, videos, prototyping, implementing proposals, reflections on process and progress, evaluation of proposals.

Following the methodology of challenge based learning, the subject is based on a coordinated set of practical works that lead to the elaboration and presentation of a proposal of technical solution of the problem posed by the actors proposing the challenge, finally contributing with some proposals of solution to solve them.

The task will be carried out in small groups of students, who can be grouped flexibly depending on the dynamics of work. The practical activity will be accompanied by a set of theoretical and methodological support sessions, as well as the tutoring of the whole learning process.

The stages of the challenge learning methodology are:

1. Discovery

The first phase involves a double recognition:

- a) on the one hand, of the field of application and the problematic of study for each one of the groups of work, coming in contact with the first actors of the subject;
- b) on the other, the tools necessary to properly carry out further research: preparation of the diagnosis, and the procedure for drawing up a Development Plan.

2. Research

Research should follow the process of analyzing the three parts of a development plan: understanding demand, supply, and needs. This stage also has a double aspect:

- a) On the one hand, the general documents of the proposed challenge are explored.
- b) On the other hand, contact is made with the necessary resources and services, conducting interviews with their managers and visiting facilities, observing users and the environment.

3. Identification of needs and proposals for improvement

The third phase is characterized by the development of a prototype for obtaining results, accurately identifying the shortcomings in each of the sectors, and coming to design proposals for improvement at an initial level for the challenge analyzed by each group.

These activities will have the reinforcement of conferences and workshops aimed at obtaining creative results.

4. Presentation of Results

Communicating the results is an essential step in the process. It will be done in three formats:

Report of the proposal
Posters
Oral presentations

Annotation: within the schedule set by the centre or degree programme, 15 minutes of one class will be reserved for students to evaluate their lecturers and their courses or modules through questionnaires.

Assessment

Continuous assessment activities

Title	Weight	Hours	ECTS	Learning outcomes
MVP reports	50	2	0.08	1, 2, 3, 4, 5
Poster and oral presentation	20	1	0.04	1, 2, 3, 4, 5
Final report	30	2	0.08	1, 2, 3, 4, 5

Continuous assessment will be used with evidence collected from each training activity of each of the phases

and in a transversal way and reflection on the learning process and progress. This subject will propose to carry out one of the evaluation items using the challenge-based learning methodology (ABR). It is about facing real and complex problems identified by professionals in order to make an appropriate definition and point out possible solutions. Challenge-based learning is today a widely recognized methodology that allows not only to deal with real case studies, but also to insert the student in environments in which they can develop their professional career in the future.

The assesment will take into account:

a) The final results of the process of elaboration of the Design Report of the proposed solution (Written report, oral presentation and poster), where the procedure and the achievement in the resolution of the posed challenge will be evaluated.

b) The gradual learning process, based on three follow-up reports (incremental MVPs).

The subject follows a continuous learning and assessment schedule that must be followed on time. The delivery dates of the tasks must be respected. Delay in deliveries will result in a penalty.

Attendance at the LAB, PROTOTYPING and Presentations sessions is mandatory.

The grade of the subject will be the average of the grades obtained in the different items evaluated. Failure to complete any of the items implies that the subject is \"Not assessable\". To be able to average, you must have obtained at least a 4 in each of the evaluable items.

In this subject, the use of Artificial Intelligence (AI) technologies is allowed as an integral part of the development of the work, provided that the final result reflects a significant contribution of the student in personal analysis and reflection. The student should clearly identify which parts have been generated with this technology, specify the tools used, and include critical reflection on how these have influenced the process and the final outcome of the activity. The non-transparency of the use of AI will be considered academic dishonesty and may lead to a penalty in the grade of the activity, or greater sanctions in cases of severity.

This subject does not provide for the single assessment system.

Recovery: Recovery requires that all items requested in the assessment have been submitted.

The follow-up exercises and the poster will be recoverable items. They may be recovered, respectively, with another examination, and with the revision of the failed poster. Due to its nature, the Written Report and the oral presentations will not be retrieved.

In the event that the student commits any irregularity that could lead to a significant variation in the grade of an assessment act, this assessment act will be graded with 0, regardless of the disciplinary process that may be instructed. In the event of several irregularities in the evaluation acts of the same subject, the final grade for this subject will be 0.

Bibliography

- Dan C. Marinescu. "Cloud Computing. Theory and Practice". Morgan-Kaufmann. 2018.

-AWS Certified Cloud Practitioner Study Guide; Ben Piper, David Clinton; Sybex (14 de junio de 2019); ISBN-10: 1119490707, ISBN-13: 978-1119490708

-Infrastructure as Code; Kief Morris; O'Reilly Media; 1 edition (June 17, 2016); ISBN-10: 1491924357, ISBN-13: 978-1491924358

-Amazon Web Services in Action, 2E; Andreas Wittig, Michael Wittig; Manning Publications; Edición: 2nd edition (30 de septiembre de 2018); ISBN-10: 1617295116, ISBN-13: 978-1617295119

-Microsoft Azure Essentials - Fundamentals of Azure, 2nd Ed; Michael Collier, Robin Shahan; 2016; https://download.microsoft.com/download/6/6/2/662DD05E-BAD7-46EF-9431-135F9BAE6332/9781509302963_I

-Mastering Cloud Computing : Foundations and Applications Programming. Buyya, Rajkumar;Vecchiola, Christian;y más Elsevier Science & Technology 2013. ISBN: ISBN number:9780124114548, ISBN number:9780124095397

This is Service Design Thinking: Basics - Tools - Cases; Marc Stickdorn; BIS Publishers; ISBN 9063692560 (ISBN13: 978906369256) 2012

Design Research: Methods and Perspectives Brenda Laurel; The MIT Press; 2003; ISBN-10 : 0262122634; ISBN-13 : 978-0262122634

Gamestorming: A Playbook for Innovators, Rulebreakers, and Changemakers; Dave Gray, Sunni Brown, James Macanufo; O'Reilly Media; ISBN-10 : 0596804172 ISBN-13 : 978-0596804176; 2010

The Lean Startup: How Today's Entrepreneurs Use Continuous Innovation to Create Radically Successful Businesses; Eric Ries ; Currency; ISBN-10 : 9780307887894, ISBN-13 : 978-0307887894; 2011

Software

Visual Studio Code

Apache Spark

Redis

Power BI

MS Power Platform

Qlik

Azure Cloud

AWS Cloud

Google Cloud

Course groups and languages

The information provided is provisional until November 30. After this date, you will be able to consult the language of each group through this [link](#). To access the information, you will need to enter the course CODE

Type of teaching	Group	Language	Semester	Shift
(TE) Theory	81	Catalan/Spanish	first semester	afternoon
(PAUL) Classroom practices	811	Catalan/Spanish	first semester	afternoon
(PLAB) Practical laboratories	811	Catalan/Spanish	first semester	afternoon
(PLAB) Practical laboratories	812	Catalan/Spanish	first semester	afternoon