

**122050058 - Màster en Analítica de dades****0058004 - Captura, preprocessat i emmagatzematge de dades****Dades generals****Codi:** 0058004**Tipus:** OB**Semestre:** 2**Crèdits:** 10.0**Llengua
d'impartició:** Català**Llengua de suport:** Català**Professorat****Professorat responsable**

Castillo Massoni, Agda (Àrea de coneixement: Ciències jurídiques i dret)

Rodó i Vila, Paula (Àrea de coneixement: Ciències tecnològiques)

Ribera Monfort, Joan (Àrea de coneixement: Ciències tecnològiques)

Informació**Presentació del mòdul:**

En iniciar un procés d'analítica de dades sovint caldrà aptar, preprocessar i emmagatzemar les dades per, posteriorment, ser analitzades.

En aquest mòdul es tracten els principals mètodes de captura de dades, com ara per exemple scripting, fluxes d'extracció, web scraping o l'accés a API, la transformació i la integració de dades provinents de diverses fonts i en diferents formats i el seu emmagatzematge en repositoris. L'objectiu final és poder validar i disposar de les dades netes que s'analitzaran. A més, també cal que l'estudiant conegui i actuï seguint els principis ètics de captació de dades i les polítiques de protecció de dades. És en aquest mòdul on s'introdueix també el que es coneix com a DataOps.

Aquesta nomenclatura, que abrevia les Data Operations, agrupa diferents grups i

persones de l'organització per aconseguir, de manera àgil i sempre intentant seguir un conjunt de bones pràctiques, les següents fites: integració de les dades, validació, gestió de les metadades i observabilitat.

Metodologia:

El mòdul està basat en una metodologia on es diferencia dos tipus d'aprenentatge que es complementen entre si. El primer està relacionat amb l'adquisició de coneixements i, el segon, amb la seva aplicació.

Aquesta metodologia combina un enfocament teòric amb un enfocament orientat a la pràctica per tal de guiar i dotar a l'estudiant amb les competències necessàries per al desenvolupament de les diferents activitats del mòdul i, més concretament, cap a la resolució del repte.

L'aplicació d'aquesta metodologia requereix una col·laboració permanent entre el professorat i l'estudiant. En conseqüència, cada setmana, aquest mòdul té associades en el calendari del curs 5 sessions lectives presencials per seminari, distribuïdes al llarg del semestre.

De forma alternativa, pels estudiants que segueixen la seva formació en modalitat virtual i, complementària pels estudiants semi-presencials, es realitzarà el seguiment dels seminaris i de les sessions de treball guiat a través del campus virtual i de les consultes i tutories en línia. A més, els estudiants que segueixen la modalitat virtual tenen la possibilitat d'assistir a les sessions presencials que considerin oportunes.

En aquestes sessions el professorat facilita als estudiants els continguts i les eines necessàries per desenvolupar les activitats i el repte del mòdul.

A més, cal afegir, el treball personal que, tot i està present durant tot el semestre, s'intensifica en les darreres setmanes per a la resolució de l'activitat final i, sobretot, del repte del mòdul.

Avaluació continuada:

Les diferents activitats d'avaluació, entre elles el repte, són fixades pel professorat i estaran disponibles en el calendari dels seminaris i del repte.

Consulteu l'apartat d'observacions.

Avaluació final:

Els estudiants que no hagin seguit o que no hagin superat l'avaluació continuada tindran la possibilitat de recuperar els seminaris pendents i el repte durant el període de recuperació previst d'acord amb el calendari del mòdul.

Consulteu l'apartat d'observacions.

Bibliografia bàsica:

Cada seminari detallarà la bibliografia bàsica.

Bibliografia complementària:

Cada seminari detallarà la bibliografia complementària.

Observacions:

El mòdul té associats uns resultats d'aprenentatge que es troben descrits en el seu pla docent. Aquests resultats d'aprenentatge són avaluats utilitzant una escala numèrica de 0 a 10 (per dècimes).

Per superar el mòdul és necessari que la qualificació de tots aquests resultats d'aprenentatge sigui superior o igual a 3 i que el valor de la mitjana aritmètica de les seves qualificacions sigui superior o igual a 5.

A més, els resultats d'aprenentatge del mòdul, cadascun en el seu nivell competencial, contribueixen a l'avaluació de les competències específiques i transversals que l'estudiant tindrà en acabar els estudis.

Competències / Resultats d'aprenentatge específics (5)

	Codi - Competències (C) / Resultat d'aprenentatge específic (RA)	Nivell
C	MAD-E003 - Dissenyar repositoris de dades per analitzar dades provinents de diferents fonts i amb diferents formats.	
RA	MAD-E003-03 - Dissenya repositoris de dades per a emmagatzemar dades heterogènies.	2
RA	MAD-E003-04 - Usa els gestors de bases de dades relacionals i no relacionals perquè les dades puguin ser analitzades.	2
C	MAD-E004 - Aplicar tècniques específiques de captura, neteja i preprocessat de dades per ser analitzades.	
RA	MAD-E004-01 - Identifica les dades essencials que cal capturar.	1
RA	MAD-E004-02 - Transforma les dades capturades en dades útils per l'anàlisi a realitzar.	1
RA	MAD-E004-03 - Aplica tècniques àgils per a la integració i validació de les dades, això com la gestió de les metadades associades.	2

Competències / Resultats d'aprenentatge transversals (1)

	Codi - Competències (C) / Resultat d'aprenentatge transversal (RA)	Àrea	Nivell
C	MAD-T004 - Actuar segons l'ètica i la responsabilitat social com a ciutadà i com a professional.	Responsabilitat personal	
RA	MAD-T004-01 - Gestiona les dades seguint criteris ètics i legals.		1

Continguts (1)

1. Pla Docent del Mòdul ([Més informació](#))