

Degree	Type	Year
2503758 Data Engineering	OT	4

Contact

Name: Pedro Enrique Dominguez Polo

Email: pedroenrique.dominguez@uab.cat

Teachers

Pedro Enrique Dominguez Polo

Teaching groups languages

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

Prerequisites

No es requereixen requisits previs per a cursar l'assignatura.

Objectives and Contextualisation

L'aparició d'Internet y de dispositius de tot tipus ha generat una gran quantitat de dades que cal processar, analitzar i visualitzar per a la presa de decisions en una empresa. Cada cop les empreses fonamenten les seves decisions a partir d'aquestes analítiques. Però,

- Sabem d'on venen les dades?.
- Les dades estan alineades a la política de l'empresa?.
- Són fiables aquestes dades?.
- Tenen la qualitat necessària?.

Per a respondre a aquestes preguntes, cal tenir un enfoc holístic respecte les dades.

L'àmbit del govern de dades permet a l'empresa administrar de forma eficaç i eficient les seves dades. Aquest àmbit ajunta processos, tecnologies i persones per adequar l'adquisició, gestió, manteniment i transformació de la informació en una empresa, per a que pugui ser compartida per tots els membres a fi de millorar la presa de decisions que es faci sobre aquesta informació.

El govern de dades és un àmbit transversal, però en l'assignatura ens centrarem en la qualitat de les dades.

Parlarem en l'assignatura de tecnologia, però també de procediments i actuacions sobre les dades dins l'empresa.

CONEIXEMENTS: Al finalitzar el curs l'alumnat ha de ser capaç de:

- Conèixer i ser capaç de proposar programes de govern de dades, de gestió de dades mestres, de qualitat de dades i/o privacitat i seguretat amb l'objectiu de millorar la gestió, la confiança, la protecció i l'explotació eficient de les dades.
- Conèixer les pràctiques que componen el govern de dades, la gestió de dades mestres, la qualitat de les dades i la seguretat i la privacitat.
- Conèixer les tècniques que componen el govern de dades, la gestió de dades mestres, la qualitat de les dades i la seguretat i la privacitat.
- Conèixer les tecnologies que participen en el govern de dades, la gestió de dades mestres, la qualitat de les dades i la seguretat i la privacitat.
- Saber analitzar l'entorn de negoci de la empresa: quines necessitat té, quins passos s'han de seguir, etc.
- Adquirir els coneixements necessaris per a seleccionar de manera adequada proveïdors i productes tenint en compta la problemàtica de l'empresa.
- Conèixer la importància de la seguretat i la privacitat de les dades en l'explotació eficient de les dades, identificant les necessitats respecte la seguretat i la privacitat de les dades.
- Conèixer les regulacions actuals que impacten el govern y l'anàlisi de les dades tant a nivell local, europeu o mundial.

Competences

- Conceive, design and implement efficient and secure data storage systems.
- Prevent and solve problems, adapt to unforeseen situations and take decisions.
- 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.
- Students must develop the necessary learning skills to undertake further training with a high degree of autonomy.

Learning Outcomes

1. Design systems that protect the privacy of personal clinical information in the field of the health sciences.
2. Prevent and solve problems, adapt to unforeseen situations and take decisions.
3. 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.
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.

Content

1. Introducció al Govern de Dades

- Què és el Govern de Dades?
- Importància del Govern de Dades: Valor de negoci
- Arquitectura de referència: Enterprise Data Management
- Aspectes legals

2. El cicle de vida de les dades

- El valor de les dades com actiu
- El cicle de vida de les dades
- Components del Govern de Dades

3. Categories d'eines pel Govern de Dades

- *Business Glossary*
- Gestió dels Metadata
- *Data Profiling*
- Qualitat de dades
- *Master Data Management*
- *Reference Data Management*
- Polítiques d'administració de la informació

4. El programa de **Data Governance**: Disseny, implantació i manteniment

- Introducció
- Àmbit
- Elements
- *Business case*

5. Criteris d'evaluació de les plataformes de **Data Governance**

- Cost total
- Mètriques de *Data Governance*
- Principals venedors d'eines
- *Business case*

6. Introducció al Govern de Dades en entorns Big Data

- *NoSQL/Hadoop*
- *Stream data*

Activities and Methodology

Title	Hours	ECTS	Learning Outcomes
Type: Directed			
Lectures	26	1.04	3, 5
Problem speaking and solving	14	0.56	1, 3, 4, 5
Project	12	0.48	1, 2, 3, 4, 5
Type: Supervised			
Practical cases	4	0.16	1, 2, 5
Type: Autonomous			
Individual study	26	1.04	3, 5
Problem solving	27	1.08	1, 3, 4, 5
Project	33	1.32	1, 2, 3, 4, 5

PLATAFORMA VIRTUAL: Campus Virtual (<http://cv.uab.cat>), aula *Govern de Dades (20-21)*.

El procés d'aprenentatge de l'alumne es fonamenta en dos tipus d'activitats presencials (Teoria i Seminaris). L'assistència a les classes de teoria i de seminaris, siguin presencials, semipresencials o online NO ÉS OBLIGATÒRIA, però SÍ MOLT RECOMANABLE.

TEORIA

Les classes de teoria s'imparteixen mitjançant classes magistrals amb transparències i pissarra. Tot el material del curs serà accessible a través de la Plataforma Virtual. Algunes de les sessions de teoria es dedicaran a resoldre exercicis plantejats a la mateixa aula amb l'objectiu d'aprofundir en els conceptes més rellevants i ensenyar a l'alumne a identificar i analitzar els errors típics.

SEMINARIS

En les sessions de SEMINARIS proposarem dos tipus d'activitats: Problemes i Projecte.

En les sessions de problemes es proporcionaran unes presentacions i uns exercicis a resoldre, que l'alumnat fa amb la tutorització del professor. Tota la documentació (enunciats scripts, resultats dels problemes) dels seminaris està disponible en la Plataforma Virtual.

En les sessions de projecte es proposarà la realització d'un projecte de govern de dades en grups de 5,6 persones. En la primera sessió es proposarà el projecte i en les sessions posteriors els alumnes treballaran en el projecte, presentant una memòria del treball i una presentació oral. Durant la realització cada grup presentarà informes e seguiment.

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

Continous Assessment Activities

Title	Weighting	Hours	ECTS	Learning Outcomes
Project Mark (portfolio and deliveries)	50%	4	0.16	1, 2, 3, 4, 5
Theory Mark (partial exams)	50%	4	0.16	1, 3, 5

L'avaluació de l'assignatura es fa de forma continuada, basada en dos indicadors: Teoria, Projecte

TEORIA

Hi haurà dos exàmens individuals per escrit, amb un pes del 50% cadascun sobre la Nota de Teoria. La primera prova (ExPar1) es farà aproximadament a la meitat del semestre i avaluarà els conceptes explicats fins a la data i la capacitat de resolució de problemes de visualització semblants als treballats en les sessions de seminaris. La segona prova (ExPar2) es realitzarà al final del semestre i avaluarà els conceptes explicats fins el moment, de teoria i la capacitat de resolució de problemes semblants als treballats en les sessions de seminaris. Ambdós exàmens s'avaluen sobre 10.

Examen de recuperació: En cas de que la nota d'algun examen parcial de teoria no arribi a l'aprovat, l'alumnat es podrà presentar a un examen de recuperació sobre els continguts de l'examen parcial suspès.

PROJECTE

La nota del projecte realitzat en sessions de seminaris té una puntuació 10. Té dos indicadors ponderats:

- Nota Primera Entrega: (nota, Prj_1) Entrega d'un informe en grup. Ponderació: 40%
- Nota Segona Entrega (nota Prj_2). Memòria i presentació oral del projecte. Ponderació: 60%.

Recuperació: S'establirà una data per a la recuperació del projecte, tant per a la primera com la segona entrega.

INDICADORS DE QUALIFICACIÓ

La nota final de l'assignatura es calcula de la següent manera (entre parèntesi quadrat les notes mínimes per fer mitja):

- $Nota Teoria = 0,5 * (Par1) + 0,5 * (Par2) [Par1, Par2 \geq 5]$
- $Nota Projecte = 0,4 * Prj_1 + 0,6 * Pr_2 [Pr_1, Pr_2 \geq 5]$

$NOTA FINAL ASSIGNATURA = 0,5 * Nota Teoria + 0,5 * Nota Projecte$

La publicació de notes es farà de forma individualitzada en la Plataforma Virtual (cv.uab.cat).

CRITERIS D'AVUACIÓ

- Per tal de poder calcular la NOTA FINAL ASSIGNATURA, caldrà superar totes les activitats avaluables (exàmens i entregues de pràctiques i projectes) amb un 5.
- NO AVALUABLE: Un alumne o alumna es considera No Avaluable (NA) únicament si no ha fet cap activitat d'avaluació. Recordem que la nota No Avaluable també corre convocatòria.
- SUSPÈS: En el cas de no arribar al mínim exigít en alguna de les activitats d'avaluació, si el càlcul de la nota final és igual o superior a 5, es posarà un 4 de nota a l'expedient.
- MATRÍCULA D'HONOR: Podrà obtenir una MH l'alumna o alumne que tingui una nota igual o superior als 9 punts, a criteri del professor. Com que el nombre de MH no pot superar el 5% de l'alumnat matriculat, es concediran a aquells o aquelles que tinguin les notes finals més altes, a criteri del professor.
- REPETIDORS: No es convalida cap part aprovada per separat (Teoria, exàmens i entregues de pràctiques) d'un curs acadèmic a un altre.

CALENDARI D'AVUACIÓ:

- Exàmens Parcial: Dates al final de la guia docent, que es publicarà en la Plataforma Virtual i que es donarà a l'alumnat el primer dia de classe.
- Exàmens de Recuperació: Segons calendari acadèmic de l'Escola d'Enginyeria.
- Exàmens i entregues de pràctiques: Dates en la normativa de pràctiques que es publicarà en la Plataforma Virtual.

Les dates d'avaluació contínua i de lliurament de treballs es publicaran a la Plataforma Virtual i poden estar subjectes a canvis de programació per motius d'adaptació a possibles incidències. Sempre s'informarà a la Plataforma Virtual sobre aquests canvis ja que s'entén que aquest és el mitjà habitual d'intercanvi d'informació entre professors i alumnat fora de l'aula.

Sense perjudici d'altres mesures disciplinàries que s'estimin oportunes, i d'acord amb la normativa acadèmica vigent, les irregularitats comeses per un alumne o alumna que puguin conduir a una variació de la qualificació es qualificaran amb un zero (0). Les activitats d'avaluació qualificades d'aquesta forma i per aquest procediment no seran recuperables. Si és necessari superar qualsevol d'aquestes activitats d'avaluació per a aprovar l'assignatura, aquesta quedarà suspesa directament, sense oportunitat de recuperar-la en el mateix curs. Aquestes irregularitats inclouen, entre d'altres:

- La còpia total o parcial d'una pràctica, informe, o qualsevol altra activitat d'avaluació
- Deixar copiar.
- Presentar un treball de grup no fet íntegrament pels membres del grup.
- Presentar com a propis materials elaborats per un tercer, encara que siguin traduccions o adaptacions, i en general treballs amb elements no originals i exclusius de l'alumna o alumne.

- Tenir dispositius de comunicació (com telèfons mòbils, *smart watches*, etc.) accessibles durant les proves d'avaluació teòric-pràctiques individuals (exàmens).

Amb les excepcions de que s'atorgarà la qualificació de "no avaluable" a l'alumnat que no participi en cap de les activitats d'avaluació, i de que la nota numèrica de l'expedient serà el valor menor entre 3,0 i la mitjana ponderada de les notes en cas que l'alumne o alumna hagi comès irregularitats en un acte d'avaluació (i per tant no serà possible l'aprovat per compensació).

EN RESUM: copiar, deixar copiar o plagiar en qualsevol de les activitats d'avaluació equival a un SUSPÈS amb nota inferior a 3,5.

Bibliography

MATERIAL DE L'ASSIGNATURA (apunts, transparències, articles): En la Plataforma Virtual.

BIBLIOGRAFIA:

- J. Ladley, *Data Governance: How to Design, Deploy and Sustain and Effective Data Governance Program*, The Morgan Kaufmann, 1st Edition, 2012, ISBN-13: 978-0124158290.
- R. S. Seiner, *Non-Invasive Data Governance*, Technics Publications; First Edition, 2014, ISBN-13: 978-1935504856.
- D. Plotkin, *Data Stewardship. An actionable Guide to Effective data management and data Governance*, Morgan Kaufmann; 1st Edition, 2013, ISBN-13: 978-0124103894.
- H. Schindlwick, *IT Governance. How to reduce and improve data Quality Through the implementation of IT Governance*, O'Reilly Media, 2019, ISBN-13: 978-1539535263.
- S. Soares, *Big Data Governance. An emerging Imperative*, CreateSpace Independent Publishing Platform, 1st Edition, 2017, ISBN-13: 978-1583473771.
- S. Sarsfield, *The Data Governance Imperative. A business strategy for corporate data*, IT Governance Publishing, 1st edition, 2016.

ENLLAÇOS WEB:

- https://www.ibm.com/uk-en/analytics/data-governance?p1=Search&p4=43700052284235969&p5=b&qclid=CjwKCAjwoZWHBhBgEiwAiM N66WbJg7T138_kJk2bdiWVzQhsbVgCB8kigKkOBICyMaA_Hm4izhmlRoCgtIQAvD_BwE&gclid=aw.ds, web d'IBM on publicita els seus recursos i productes que ofereix sobre Data Governance. (Darrer accés: juliol 2021).
- <http://library.ahima.org/doc?oid=301363#.YOXIOkztZaQ>, Information Governance: Bibliography of Articles (2016 update). (Darrer accés: juliol 2021).
- https://www.collibra.com/blog/what-is-data-governance?utm_source=Google-Ads&utm_medium=Paid&utm_campaign=EU-BL-Search-NB-General-Data-Governance-TOFU&utm_content=What-is-a-dat-66avDgTi8gC4HYgfejkbZbSR0saMDnQbMMBghWGgbRXwQmUlf3IfUvxoCDCoQAvD_BwE, web que publica els productes de Data Governance de l'empresa Collibra. (Darrer accés: juliol 2021).

Software

En l'assignatura es treballarà amb el software IBM InfoSphere Information Server for Data Quality, amb llicència acadèmica que es sol·licitarà a IBM.

Sempre que sigui possible es buscarà posar a disposició dels alumnes software lliure o amb llicència educativa.

Language list

Name	Group	Language	Semester	Turn
(PAUL) Classroom practices	1	Catalan	second semester	morning-mixed

PROVISIONAL