

Titulació	Tipus	Curs
4318290 Arxivística i Governança de la Informació	OB	2

Professor/a de contacte

Nom: Eloi Puertas Prats

Correu electrònic: eloi.puertas@uab.cat

Equip docent

Jordi Serra Serra

Idiomes dels grups

Podeu consultar aquesta informació al [final](#) del document.

Prerequisits

1. Haver cursat l'assignatura "A04. Sistemas de información y arquitectura de sistemas"
2. Haver cursat l'assignatura "A09. Descripción y recuperación de la información"

Objectius

1. Conèixer el cicle de vida dels dades i la seva gestió.
2. Conèixer el context de producció de les dades.
3. Aplicar els principis arxivístics a la gestió de dades.
4. Conèixer i comprendre les principals eines i sistemes de gestió de dades
5. Conèixer sistemes gestors de dades i bases de dades
6. Conèixer models de governança de dades, normes i estàndards
7. Conèixer i comprendre els sistemes bàsics d'ús, explotació i visualització de dades.

Resultats d'aprenentatge

1. CA21 (Competència) Establir les dades de qualitat d'una organització.
2. CA22 (Competència) Dissenyar els criteris i els formats per a la gestió del cicle de vida de les dades d'una organització.
3. KA30 (Coneixement) Descriure el cicle de vida de les dades.
4. KA31 (Coneixement) Identificar tipus i formats de dades.

5. KA32 (Coneixement) Reconèixer els sistemes per a la governança de dades: repositoris, plataformes d'arquitectura de dades i sistemes de bases de dades.
6. SA23 (Habilitat) Utilitzar les eines i els sistemes de gestió de dades principals.
7. SA24 (Habilitat) Aplicar tècniques per a la governança de dades en les organitzacions.
8. SA25 (Habilitat) Aplicar els principis arxivístics a la gestió de dades.

Continguts

- 1.1. Les dades a les organitzacions (introducció)
- 1.2. On es produeixen les dades?
 - 1.2.1. Formes de captura i generació de dades (tràmits, sensors, etc.)
 - 1.2.2. Els models per estructurar les dades (mestres, referencials, etc.)
 - 1.2.3. Arquitectures per a l'emmagatzemament (tipus de bases de dades)
 - 1.2.3.1. Ús de DB relacionals (SQL)
 - 1.2.3.2. Ús de DB NoSQL (Hadoop i HDFS, MongoDB, etc.)
- 1.3. Com s'utilitzen les dades?
 - 1.3.1. La preparació de les dades
 - 1.3.1.1. Neteja de les dades (data cleansing)
 - 1.3.1.2. Preparació per a l'explotació (cubs, BI, etc.)
 - 1.3.1.3. Consolidació del tractament (ETL, RPA, etc.)
 - 1.3.1.4. Formats de les dades a depurar (CSV, JSON, XML, etc.)
 - 1.3.1.5. Neteja i preparació de dades amb Python
 - 1.3.2. L'explotació i ús de les dades
 - 1.3.2.1. Visualització de dades
 - 1.3.2.2. Anàlisi avançada estadística, o basada en ML i IA
 - 1.3.2.3. Aplicació pràctica d'algoritmes d'anàlisi avançada
- 1.4. El govern integrat de la dada
 - 1.4.1. Identificació i catalogació de dades
 - 1.4.2. El control del llinatge de la dada
 - 1.4.3. Virtualització de l'accés a les dades
 - 1.4.4. Aspectes legals i de seguretat
 - 1.4.5. Els vincles amb l'arxivística

Activitats formatives i Metodologia

Títol	Hores	ECTS	Resultats d'aprenentatge
Tipus: Dirigides			
Sessions teòriques	45	1,8	CA21, CA22, KA30, KA31, KA32, SA23, SA24, SA25, CA21
Tipus: Supervisades			
Exercici 1: neteja, depuració i preparació d'un dataset.	30	1,2	CA21, CA22, KA30, KA31, KA32, SA23, SA24, SA25, CA21
Exercici 2: creació d'una visualització de dades simple.	30	1,2	CA21, CA22, KA30, KA31, KA32, SA23, SA24, SA25, CA21
Exercici 3: execució d'una anàlisi avançada sobre un conjunt de dades (regressió o anàlisi de conglomerats).	20	0,8	CA21, CA22, KA30, KA31, KA32, SA23, SA24, SA25, CA21
Tipus: Autònomes			
Lectura de Materials	90	3,6	CA21, CA22, KA30, KA31, KA32, SA23, SA24, SA25, CA21
Prova final: test de coneixements generals de l'assignatura.	10	0,4	CA21, CA22, KA30, KA31, KA32, SA23, SA24, SA25, CA21

Les activitats formatives autònomes seran lectura de materials i la preparació per a la prova final de test de coneixements generals de l'assignatura.

Les dirigides seran sessions teòriques magistrals.

Les supervisades seran 3 exercicis pràctics per fer a casa amb les explicacions rebudes a classe.

Nota: es reservaran 15 minuts d'una classe, dins del calendari establert pel centre/titulació, per a la complementació per part de l'alumnat de les enquestes d'avaluació de l'actuació del professorat i d'avaluació de l'assignatura/mòdul.

Avaluació

Activitats d'avaluació continuada

Títol	Pes	Hores	ECTS	Resultats d'aprenentatge
Exercici 1: neteja, depuració i preparació d'un dataset.	25% de la nota final	0	0	CA21, SA24

Exercici 2: creació d'una visualització de dades simple.	25% de la nota final	0	0	KA31, SA23, SA24
Exercici 3: execució d'una anàlisi avançada sobre un conjunt de dades (regressió o anàlisi de conglomerats).	20% de la nota final	0	0	CA21, KA32, SA23
Prova final: test de coneixements generals de l'assignatura.	30% de la nota final	0	0	CA21, CA22, KA30, KA31, KA32, SA23, SA24, SA25

Tant l'exercici 1 com el 2 valdrà un 25 de la nota final. El 3er valdrà un 20% i l'examen final el 30%

Bibliografia

- Benfeldt, O., Persson, J. S., & Madsen, S. (2020). Data Governance as a Collective Action Problem. *Information Systems Frontiers*, 22(2), 299-313. <https://doi.org/10.1007/s10796-019-09923-z>
- Earley, S., & Henderson, D. (Ed.). (2017). *DAMA-DMBOK: Data management body of knowledge* (2nd edition). Data Management Association.
- Ghavami, P. (2020). *Big data management: Data governance principles for big data analytics* (1a ed.). De Gruyter.
- Khatri, V., & Brown, C. V. (2010). Designing data governance. *Communications of the ACM*, 53(1), 148-152.
- Laurent, A., Laurent, D., & Madera, C. (Ed.). (2019). *Data lakes*. ISTE Ltd / John Wiley and Sons Inc.
- Lemieux, V. L., Gormly, B., & Rowledge, L. (2014). Meeting Big Data challenges with visual analytics: The role of records management. *Records Management Journal*, 24(2), 122-141.

Programari

- Microsoft Power BI (desktop version)
- Microsoft Sharepoint
- Anaconda (Python)

Llista d'idiomes

Nom	Grup	Idioma	Semestre	Torn
(TE) Teoria	1	Català	primer quadrimestre	tarda