

**FITXA IDENTIFICATIVA****Dades de l'Assignatura**

Codi	34695
Nom	Sistemes d'informació de nova generació
Cicle	Grau
Crèdits ECTS	6.0
Curs acadèmic	2022 - 2023

Titulació/titulacions

Titulació	Centre	Curs	Període
1400 - Grau d'Enginyer Informàtica	Escola Tècnica Superior d'Enginyeria	4	Segon quadrimestre

Matèries

Titulació	Matèria	Caràcter
1400 - Grau d'Enginyer Informàtica	16 - Matèria Optativa	Optativa

Coordinació

Nom	Departament
FUERTES SEDER, ARIADNA	240 - Informàtica
VES CUENCA, ESTHER DE	240 - Informàtica

RESUM

L'enorme volum de dades que s'emmagatzemen avui en dia està creixent de forma exponencial. L'any 2000 la informació emmagatzemada era de l'ordre dels 800.000 petabytes (PB). S'estima que aquesta xifra arribarà als 35 zetabytes (ZB) el 2020. Només Twitter genera més de 7 terabytes (TB) de dades cada dia, Facebook 10 TB i algunes empreses generen terabytes de dades cada hora de cada dia de l'any. Ja no és estrany trobar empreses que compten amb capacitats d'emmagatzematge de dades de l'ordre del petabyte. Els motors d'aquest creixement són principalment el desenvolupament de la web, el nombre creixent de proveïdors de xarxes socials, el desenvolupament dels serveis en el núvol i en general el model de "programari com a servei".

Paral·lelament, la capacitat per a tractar i explorar tota aquesta sobrecàrrega d'informació i aplicar les tècniques d'extracció de significat útil són crítiques en qualsevol organització, empresa i institució actual. No és estrany, per tant, que la demanda de sistemes d'alta escalabilitat haja crescut també de manera exponencial per adaptar-se al volum de dades i a les necessitats de processament.



Tot i que la bases de dades relacionals han demostrat la seva capacitat per adaptar-se amb èxit a la creixent necessitat d'emmagatzematge, ara més que mai calen nous paradigmes de persistència sòlids, fiables i escalables capaços d'atendre les demandes dels nous sistemes d'informació. Són dues les estratègies que s'estan adoptant per resoldre aquest problema:

1. 1. D'una banda, els sistemes relacionals tradicionals es poden adaptar per respondre a les necessitats d'escalabilitat mitjançant la paral·lelització (major velocitat), la distribució (major capacitat d'emmagatzematge i procés) i la federació i replicació (escalabilitat horitzontal mitjançant la simple addició de nodes).
2. 2. D'altra banda, la introducció de nous paradigmes de bases de dades, com ara NoSQL o Big Data, d'intentar atacar el problema proposant noves estructures d'emmagatzematge més versàtils, encara que siga a costa de la pèrdua de certes funcionalitats tradicionalment considerades com a bones.

En aquesta assignatura s'estudien i analitzen els avantatges i inconvenients de les dues estratègies i es presenten les noves tecnologies de persistència que el professional de les Tecnologies de la Informació i la Comunicació (TIC) ha necessàriament conèixer per integrar-se al mercat laboral actual.

CONEIXEMENTS PREVIS

Relació amb altres assignatures de la mateixa titulació

No heu especificat les restriccions de matrícula amb altres assignatures del pla d'estudis.

Altres tipus de requisits

Per a una adequada comprensió de l'assignatura és necessari haver cursat les assignatures Bases de Dades de 2n curs i Sistemes de Gestió de Bases de Dades de 3r curs.

COMPETÈNCIES

1400 - Grau d'Enginyer Informàtica

- SI1 - Capacitat per integrar solucions de tecnologies de la informació i la comunicació i processos empresarials per satisfer les necessitats d'informació de les organitzacions, permetent-los aconseguir els objectius d'una forma efectiva i eficient, i donar-los així avantatges competitius.
- SI3 - Capacitat per participar activament en l'especificació, el disseny, la implementació i el manteniment dels sistemes d'informació i comunicació.



RESULTATS DE L'APRENTATGE

- • Comprendre la motivació dels sistemes de gestió de bases de dades paral·leles, distribuïts i federats.
- • Disposar dels fonaments bàsics per analitzar les diferents arquitectures, les implicacions en el rendiment, la velocitat, l'escalabilitat i la seguretat de les diferents estratègies de particions de les dades i avaluar i optimitzar les consultes sobre dades distribuïts.
- • Comprendre què és data warehousing i per què és important en el suport per a la presa de decisions.
- • Comprendre el model de dades multidimensional i el tipus d'anàlisi de dades que facilita.
- • Comprendre què és data mining i disposar dels fonaments bàsics de les tècniques utilitzades en mineria de dades.
- • Conèixer l'algoritme MapReduce i la seva utilització en l'anàlisi ràpida de quantitats immenses de dades.

Conèixer els diferents sistemes d'emmagatzematge massiu d'informació coneguts genèricament com a sistemes NoSQL i disposar dels fonaments bàsics per definir, dissenyar i implementar sistemes d'informació utilitzant aquests sistemes.

DESCRIPCIÓ DE CONTINGUTS

1. Sistemes avançats de gestió de bases de dades

Bases de dades distribuïdes.
Bases de dades paral·leles.
Federació de bases de dades.
Replicació.

2. Sistemes de suport per a la presa de decisions

Data warehouses
Data mining

3. Big data i NoSQL

Terminologia
Bases de dades NoSQL
MapReduce
Emmagatzematge distribuït.

**4. Recuperació d'informació i dades XML**

Recuperació d'informació.

Model de dades per a XML.

XQuery.

Avaluació eficient de consultes XML.

VOLUM DE TREBALL

ACTIVITAT	Hores	% Presencial
Classes de teoria	30,00	100
Pràctiques en laboratori	20,00	100
Pràctiques en aula	10,00	100
Elaboració de treballs en grup	5,00	0
Elaboració de treballs individuals	20,00	0
Estudi i treball autònom	10,00	0
Lectures de material complementari	5,00	0
Preparació d'activitats d'avaluació	15,00	0
Preparació de classes de teoria	10,00	0
Preparació de classes pràctiques i de problemes	25,00	0
TOTAL	150,00	

METODOLOGIA DOCENT

La docència consistirà en una combinació de lliçons teòriques, sessions de problemes i activitats de caràcter pràctic a realitzar per part de l'estudiant. Esta docència quedarà complementada amb el treball personal de l'alumnat, centrat en l'estudi, en la resolució de problemes, i en la preparació de treballs per a entregar. A més, es realitzaran sessions de laboratori basades en treball amb ordinador.

- Les activitats teòriques consistiran en la realització de classes d'una hora i mitja de duració en què es desenrotllaran els temes proporcionant una visió global i integradora, analitzant amb major detall els aspectes clau i de major complexitat, fomentant, en tot moment, la participació de l'alumnat.
- Les activitats pràctiques consistiran en la realització de sessions de problemes i qüestions en l'aula d'hora i mitja de duració, així com la realització de seminaris, en els que s'abordaran dos temes sobre aplicacions i aspectes menys formals de l'assignatura. Complementen les lliçons expositives amb l'objectiu d'aplicar els conceptes bàsics i ampliar-los amb el coneixement i l'experiència que vagen adquirint durant la realització dels treballs proposats. Com a exemples d'estes activitats podem mencionar les següents:
 - o Classes de problemes i qüestions en aula
 - o Seminaris aplicats en pissarra i aula d'ordinadors
 - o Sessions de discussió i resolució de problemes i exercicis prèviament treballats pels



- estudiants
 - o Tutories individualitzades
- Les sessions de laboratori consistiran en la resolució de problemes relacionats amb els continguts teòrics per mitjà de programes informàtics en sessions de dos hores i mitja de duració
- El treball personal de l'alumnat intenta fomentar el treball autònom i consistirà, fonamentalment, en tres aspectes:
 - o La preparació de les classes amb antelació i la lectura de textos recomanats i preparació d'exàmens (estudi).
 - o La resolució de problemes proposats pel professor o L'elaboració de treballs que s'entregaran al professor.
 - o La preparació de les sessions de laboratori amb anterioritat a l'assistència a l'aula de pràctiques informàtiques.
- El Treball en xicotets grups d'estudiants (2-4) per a la realització d'activitats, qüestions, problemes fora de l'aula. Esta tasca complementa el treball individual i fomenta la capacitat d'integració en grups de treball.

S'utilitzarà la plataforma d'e-learning (Aula Virtual) de la Universitat de València com a suport de comunicació amb els estudiants. A través d'ella es tindrà accés al material didàctic utilitzat en classe, així com els problemes i exercicis a resoldre.

AVALUACIÓ

El treball que s'espera que els alumnes realitzen ho podem classificar en 2 tipus:

1. Treball autònom d'autoaprenentatge.
2. Treball supervisat.

1. El treball autònom d'autoaprenentatge consisteix en les activitats que l'alumne realitza fora de l'horari de docència, de manera independent o dirigida pel professor, i que està orientat a adquirir coneixements, capacitats i destreses de forma autodidacta. Principalment seran activitats que el professor indique que s'han de realitzar durant el curs (lectures recomanades, realització de problemes, investigació de certs temes, etc.) però que no seran avaluades directament (qualificades) pel professor, encara que tindran incidència en altres activitats que sí que seran avaluades i qualificades pel professor, com són el cas dels debats i presentacions realitzats durant les hores de docència o la realització d'exàmens.

2. El treball supervisat consisteix en totes les activitats que realitzarà l'alumne a petició del professor i que este monitoritzarà per a avaluar el grau de superació de l'assignatura. Estos treballs seran de 3 tipus:

- a) Problemes o activitats Individuals
- b) En grup



c) Pràctiques d'aula d'informàtica, d'assistència obligatòria.

Les característiques d'estos treballs són:

- • Han de ser avaluades pel professor.
- • Han de tindre data d'entrega, y/o realitzar-se de manera presencial en el moment de ser plantejades, la qual cosa obligaria a l'assistència en els dits casos.
- • Són de realització obligatòria per part de l'alumne per a l'aplicació de l'avaluació contínua.

En primera convocatòria, l'alumne realitzarà al llarg del curs diverses proves objectives individuals durant les hores lectives i que consistiran en tant qüestions teoricopràctiques com problemes o debats (N_Pruebas) que no seran recuperables. A més, caldrà entregar distintes activitats corresponents a les activitats exercides en el laboratori (N_Actividades). Existirà finalment un examen (N_Examen).

La qualificació final, en primera convocatòria, s'obtindrà per mitjà de la següent ponderació de les distintes activitats i proves:

$\text{Nota Final} = 20\% \text{ N_Examen} + 40\% \text{ N_Pruebas} + 40\% \text{ N_Actividades}$

En segona convocatòria la qualificació final s'obtindrà basant-se en la fórmula:

$\text{Nota Final} = 60\% \text{ N_Examen} + 40\% \text{ N_Actividades}$

Serà necessari obtindre, almenys, 5 sobre 10 en cada una de les parts per a poder mediar la nota.

En qualsevol cas, l'avaluació de l'assignatura es farà d'acord amb el Reglament d'avaluació i qualificació de la Universitat de València per a títols de grau i de màster, aprovat en la sessió del Consell de Govern de 30 de maig de 2017. (ACGUV 108/2017).

Esta avaluació part de la premissa que la docència en la Universitat de València és, per definició, una docència presencial. En este sentit, l'alumne ha de tindre present que l'assistència, tant a les classes teòriques com a aquelles de caràcter pràctic, és fonamental per a un adequat seguiment dels continguts de l'assignatura. L'alumne ha de tindre present igualment la possibilitat d'una matrícula a temps parcial, excepte en el cas dels alumnes de 1r, quan no li siga possible assistir a la totalitat de les assignatures que componen un curs complet (60 crèdits). Amb tot, s'establirà la possibilitat, en els casos que estiguen adequadament justificats i per a aquells alumnes que ho sol·liciten, la possibilitat de ser avaluat sense necessitat d'assistir a la totalitat o a part de les classes. En estos casos l'alumne ha de procedir de la manera següent:

- • S'ha de comunicar al principi del curs al professor/es responsable/s de l'assignatura la incidència per la qual li és impossible assistir a classe, que ha d'estar adequadament justificada de forma documental
- • El professor responsable, a la vista d'esta informació decidirà la possibilitat d'avaluació sense assistència total o parcial a les classes de l'assignatura.



Els alumnes que es troben en esta situació, hauran de presentar, per a ser avaluats, la totalitat de treballs requerits pel professor (no necessàriament idèntics als requerits durant el curso) així com també podran ser cridats a defensar-los oralment davant del propi professor, i realitzaran una prova de coneixements adquirits. El pes dels treballs en la qualificació final serà d'un 50% i el de la prova de coneixements el 50% restant.

REFERÈNCIES

Bàsiques

- Ramez A. Elmasri, Shmkant B. Navathe Fundamentos de Sistemas de Bases de Datos. Addison Wesley, 2002.
- R. Ramakrishnan, J. Gehrke Database Management Systems. McGraw-Hill, 2000.
- S. Tiwari. Professional NoSQL. Wrox, 2011.
- T. White. Hadoop: the definitive guide. Wrox, 2009

Complementàries

- Referencia c1: Priscilla Walmsley. XQuery. O'Reilly Media, Inc., 2007
- H. Inmon. Building the Data Warehouse, John Wiley & Sons, 2005
- D. Miner. MapReduce design patterns. O'Reilly 2012
- Christof Strauch. NoSQL databases. <http://www.christof-strauch.de/nosql dbs.pdf>
- C. Adamson. Mastering data warehouse aggregates. Willey, 2006