



Pla Docent de l'Assignatura

Guia Docent

Nom de l'assignatura: Tecnologies de la Informació

Curs acadèmic: 2012-2013

Curs: Segon

Trimestre: Tercer

Estudis: Grau en Enginyeria en Informàtica (Obligatòria), Grau en Enginyeria en Telemàtica (Optativa) i Grau en Enginyeria en Sistemes Audiovisuals (Optativa)

Codi assignatura: 21419-21771-21658

Nombre de crèdits ECTS: 4

Nombre total d'hores de dedicació: 100 hores

Llengua o llengües de docència: Català i Castellà

Professorat: Ricardo Baeza, Georgina Ramírez, Ruth García, Eduardo Graells

Professorat responsable: Ricardo Baeza

1. Dades descriptives de l'assignatura

Nom de l'assignatura: Tecnologies de la Informació

Curs acadèmic: 2012-2013

Curs: Segon

Trimestre: Tercer

Estudis: Grau en Enginyeria en Informàtica (Obligatòria), Grau en Enginyeria en Telemàtica (Optativa) i Grau en Enginyeria en Sistemes Audiovisuals (Optativa)

Codi assignatura: 21419-21771-21658

Nombre de crèdits ECTS: 4

Nombre total d'hores de dedicació: 100 hores

Llengua o llengües de docència: Català i Castellà

Professorat: Ricardo Baeza, Georgina Ramírez, Ruth García, Eduardo Graells

Professorat responsable: Ricardo Baeza

2. Presentació de l'assignatura

Aquesta assignatura pretén introduir els conceptes bàsics de recuperació de la informació i la seva aplicació a la tecnologia dels cercadors Web. Es donarà èmfasis als models de recuperació i a les tècniques d'avaluació més importants actualment, incloent-hi tècniques de jerarquització de pàgines Web. S'introduirà l'índex de text més utilitzat, l'índex invertit, com construir-lo i com utilitzar-lo. S'estudiarà també com resoldre problemes d'escalabilitat quan el volum de dades i de les consultes augmenta. Per acabar, s'aplicaran aquests conceptes a altres tipus de dades com el text estructurat en XML i multimèdia, a més de la seva aplicació a dominis específics, com bases documentals legals o de salut.

Prerequisits per al seguiment de l'itinerari formatiu: és necessari tenir coneixements tant de programació com d'estructures de dades bàsiques. Es recomana conèixer també operacions relacionades amb el processament de cadenes de caràcters (strings), a més de nocions d'XML, algorítmica i aprenentatge automàtic. Es pressuposa a més coneixement avançat en l'ús de cercadors Web.

3. Competències a assolir a l'assignatura

Competències¹ a treballar a l'assignatura segons l'indicat en el pla d'estudis del grau.

Competències transversals	Competències específiques
<i>Instrumentals</i> G1. Capacitat d'anàlisi i síntesi G2. Capacitat d'organització i planificació G3. Capacitat per aplicar els coneixements a l'anàlisi de situacions i la resolució de problemes G4. Habilitat en la cerca i la gestió de la informació G5. Habilitat en la presa de decisions G6. Capacitat de comunicar-se amb propietat de forma oral i escrita en català i en castellà, tant davant audiències expertes com a inexpertes.	<i>Competències Específiques Professionals</i> H4. Aprendre de manera autònoma nous coneixements i tècniques adequats per a la concepció, el desenvolupament o l'explotació de sistemes informàtics. <i>Competències Específiques de Formació Bàsica</i> B14-T. Conèixer els fonaments teòrics de la programació i utilitzar de forma pràctica els mètodes i llenguatges de programació per al desenvolupament de sistemes programari. B16-T. Conèixer els conceptes bàsics de l'arquitectura d'ordinadors i dels servidors, així com els principis dels sistemes operatius.

¹ **Competències transversals:** Aquelles que es requereixen en l'exercici de qualsevol titulació o carrera (comunicació verbal i escrita, pensament analític i sistèmic, resolució de problemes, creativitat, etc.). Es classifiquen en:

- **Instrumentals:** Inclouen habilitats cognitives, metodològiques, tecnològiques i lingüístiques. (Ex: capacitat d'organitzar i planificar, capacitat de comunicar-se amb propietat de forma oral i escrita en català, castellà i/o anglès, tant davant d'audiències expertes com inexpertes).
- **Interpersonals:** Es defineixen com habilitats que tendeixen a facilitar els processos d'interacció social i cooperació. (Ex: capacitat per a treballar en grup, expressió del compromís ètic/social).
- **Sistèmiques o integradores:** Suposen una combinació de comprensió, sensibilitat i coneixement que permeten veure com s'agrupen i s'estableixen relacions entre les parts d'un tot. Aquestes competències requereixen, com a base, l'adquisició prèvia de competències instrumentals i interpersonals. (Ex: capacitat d'adaptar-se a nous contextos d'aprenentatge)

Competències específiques: Es relacionen amb els coneixements i pràctiques concretes del grau. (Ex: capacitat de descriure, programar, validar i optimitzar protocols i interfaces de comunicació en els diferents nivell d'una arquitectura de xarxes)

<i>Interpersonals</i> G8. Capacitat de treball en equip <i>Sistèmiques</i> G14. Capacitat de motivació per la qualitat i per l'assoliment	<i>Competències Específiques d'Enginyeria en Informàtica</i> IN15. Conèixer els fonaments de les bases de dades no estructurades (documentals i multimèdia) i les tècniques relacionades de classificació i recuperació de la informació. IN36. Conèixer i comprendre els principis de les diferents modalitats i arquitectures d'informació multimèdia, i ser capaç d'aplicar les més adequades per a cada problema. <i>Competències Específiques Comunes a la branca de Telecomunicació</i> T7. Capacitat de conèixer i utilitzar l'arquitectura i metodologia de disseny, verificació i validació de programari. T8. Capacitat de realitzar programació en temps real, concurrent, distribuïda i basada en esdeveniments, així com el disseny d'interfícies persona-computador. <i>Competències de tecnologia específica: Sistemes Audiovisuals</i> AU37. Conèixer les tècniques bàsiques de la mineria de dades i textos en la web i capacitat per aplicar-les per a problemes concrets. AU39. Domini de les tècniques avançades de la cerca intel·ligent d'informació en la web.
---	--

4. Continguts

Blocs de contingut

1. Conceptes de recuperació de la informació
2. Índexs de cerca en text
3. Cercadors Web
4. Aplicacions de les tecnologies de cerca

Bloc de contingut 1. – Conceptes de recuperació de la informació

Conceptes	Procediments	Actituds
1. Models 2. Processament del text i la consulta 3. Avaluació	1. Realització de programes	1. Claredat i pulcritud en la realització de les pràctiques

Bloc de contingut 2. - Índexs

Conceptes	Procediments	Actituds
1. Construcció 2. Cerca	1. Realització de programes	1. Claredat i pulcritud en la realització de les pràctiques

Bloc de contingut 3. – Cercadors Web

Conceptes	Procediments	Actituds
1. Arquitectura d'un cercador Web 2. Jerarquització de pàgines Web 3. Escalabilitat	1. Realització de programes 2. Ús avançat de cercadors Web	1. Claredat i pulcritud en la realització de les pràctiques

Bloc de contingut 4. – Aplicacions de les tecnologies de cerca

Conceptes	Procediments	Actituds
1. Multimèdia 2. Text estructurat en XML 3. Dominis específics	1. Resolució de problemes a classe	1. Participació activa en els seminaris

5. Avaluació del nivell d'assoliment de les competències

Per superar l'assignatura cal aprovar l'examen de final de curs i realitzar 5 pràctiques de laboratori. Aquestes pràctiques seran revisades i puntuades pels professors durant les sessions de laboratori, i la puntuació només serà igual o superior a 5 si s'han fet de forma profitosa. Les pràctiques es realitzaran en grups de 3 alumnes.

En el cas que un grup no hagi pogut lliurar alguna de les pràctiques, l'avaluació es realitzarà en una entrevista personal amb el professor de laboratori, que cal concertar en hores de tutoria abans del període d'exàmens del trimestre.

Al final del trimestre es realitzarà un examen que inclourà la matèria corresponent de tot el trimestre, i cal treure un mínim de 5.

La nota final de l'assignatura és la suma ponderada de tres parts:

- Seminaris, 10%, no recuperable.
- Pràctiques, 30%, no recuperable.
- Exàmen final, 60%, recuperable al juliol.

Per aprovar l'assignatura, cal aprovar cada una de les tres parts per separat.

6. Bibliografia i recursos didàctics

6.1 Fonts d'informació per a l'aprenentatge. Bibliografia bàsica (suport paper i parcialment digital)

- BAEZA-YATES, Ricardo; RIBEIRO-NETO, Berthier: Modern Information Retrieval, segona edició. Addison-Wesley, 2011. Pàgina Web: www.mir2ed.org.

6.2 Fonts d'informació per a l'aprenentatge. Bibliografia complementària (digital)

- MANNING, Chris; RAGHAVAN, Prabhakar; SCHÜTZE, Heinrich: Introduction to Information Retrieval, Cambridge University Press. 2008. Pàgina Web: www.informationretrieval.org

7. Metodologia

7.1 Enfocament metodològic de l'assignatura

A les sessions de teoria, totes en grup gran, s'introduiran els conceptes teòrics bàsics i es mostraran els procediments adequats per a la resolució dels problemes. A les sessions de seminari es discutiran les aplicacions dels conceptes teòrics introduïts. A les sessions de laboratori es realitzaran pràctiques de programació de diferents elements d'un cercador. L'objectiu és doble: per un cantó han de servir per entendre i consolidar els conceptes teòrics i per l'altre serveixen com a indicadors d'avaluació de l'assoliment de les competències relacionades amb les tecnologies de cerca.

El treball fora de l'aula consistirà bàsicament en la recerca d'informació complementària, la resolució de problemes proposats, la preparació de les pràctiques i la realització d'estudis previs.

7.2 Organització temporal: sessions, activitats d'aprenentatge i temps estimat de dedicació

Les sessions presencials a l'aula s'organitzen així:

Bloc de continguts	Gran grup	Laboratori	Seminari
Introducció	T ₁		
1. Conceptes de recuperació de la informació	T ₁ T ₂ T ₃ T ₄	P ₂	
2. Índexs de cerca en text	T ₅ T ₆	P ₃ , P ₄	
3. Cercadors Web	T ₇ T ₈ T ₉	P ₁ , P ₅	S ₄
4. Aplicacions			S ₁ S ₂ S ₃

Els lliuraments previstos seran a les sessions de laboratori i a les sessions de seminari.

Les hores estimades de dedicació són:

	Activitats dins de l'aula	Activitats fora de l'aula	Avaluació			
	Gran grup	Laboratori	Seminari	Preparació de pràctiques	Estudi personal i realització de problemes	Examen
Introducció	2					
1. Conceptes de recuperació de la informació	6	2		4	12	
2. Índexs	4	4		8	8	
3. Cercadors Web	6	4	2	8	12	
4. Aplicacions			6		9	
Avaluació						3
Total	18	10	8	20	41	3

Total: 100

8. Programació d'activitats

- Programació de sessions presencials

	Dilluns 16:30-18:30	Dimecres 18:30-20:30	Divendres 14:30-16:30
Setmana 1 (8 - 12 abril)	Introducció a la recuperació de la informació (RI)	Cercadors en la pràctica: Terrier i Lucene (P101)	Cercadors en la pràctica: Terrier i Lucene (P102)
Setmana 2 (15 - 19 abril)	Models de recuperació de la informació	Avaluació en RI	Processament del text i de la consulta
Setmana 3 (22 - 26 abril)	Índexs invertits: construcció	Avaluació en RI (P102)	Índexs invertits: cerca
Setmana 4 (29 abril-3 maig)	Avaluació en RI (P101)	FESTIU	RI en dades XML (S103) Aplicacions (S102)
Setmana 5 (6 - 10 maig)	RI en dades XML (S102) Aplicacions (S103)		RI en dades XML (S101) Aplicacions (S104)
Setmana 6 (13 - 17 maig)	RI en dades XML (S104) Aplicacions (S101)	Construcció de índexs invertits (P101)	Construcció de índexs invertits (P102)
Setmana 7 (20 - 24 maig)	FESTIU	Cerca en índexs invertits (P102)	Cerca en índexs invertits (P101)
Setmana 8 (27 maig- 31 maig)	Cercadors Web	Implementació de un Cercador (P101)	Jerarquització de resultats en la Web
Setmana 9 (3 - 7 juny)	Escalabilitat en sistemes RI	Implementació de un Cercador (P102)	RI Multimèdia (S103) Jerarquització Web amb AA (S104)
Setmana 10 (10 -14 juny)	RI Multimèdia (S104) Jerarquització Web amb AA (S103)	RI Multimèdia (S102) Jerarquització Web amb AA (S101)	RI Multimèdia (S101) Jerarquització Web amb AA (S102)
Setmana 11 (17 -12 juny)	EXAMENS	EXAMENS	EXAMENS

- Llistat d'activitats (amb termini de lliurament i avaluació)

Activitat	Data enunciat	Data lliurament	Data de lliurament de resultats
Practica	16/4	23/4	30/4
Practica	25/4	2/5	9/5
Problema Integrador	4/5	11/5	18/5
Problema Integrador	4/5	11/5	18/5
Practica	16/5	23/5	30/5
Problema Integrador	1/6	8/6	15/6
Practica	4/6	11/6	18/6
Problema Integrador	8/6	15/6	18/6