



# Pla Docent de l'Assignatura

## Guia Docent

**Nom de l'assignatura:** Tecnologies de la Informació

**Curs acadèmic:** 2011-2012

**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, David Nettleton, Georgina Ramírez

**Professorat responsable:** Ricardo Baeza

## **1. Dades descriptives de l'assignatura**

**Nom de l'assignatura:** Tecnologies de la Informació

**Curs acadèmic:** 2011-2012

**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, David Nettleton, Georgina Ramírez

**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

Competències<sup>1</sup> a treballar a l'assignatura segons l'indicat en el pla d'estudis del grau.

| Competències transversals                                                                                                                                                                                                                                                                                                                                                                                                                                                                  | Competències específiques                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <i>Instrumentals</i><br><br>G1. Capacitat d'anàlisi i síntesi<br><br>G2. Capacitat d'organització i planificació<br><br>G3. Capacitat per aplicar els coneixements a l'anàlisi de situacions i la resolució de problemes<br><br>G4. Habilitat en la cerca i la gestió de la informació<br><br>G5. Habilitat en la presa de decisions<br><br>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><br><br>H4. Aprendre de manera autònoma nous coneixements i tècniques adequats per a la concepció, el desenvolupament o l'explotació de sistemes informàtics.<br><br><i>Competències Específiques de Formació Bàsica</i><br><br>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.<br><br>B16-T. Conèixer els conceptes bàsics de l'arquitectura d'ordinadors i dels servidors, així com els principis dels sistemes operatius. |

<sup>1</sup>**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)

|                                                                                                                                                                       |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <p><i>Interpersonals</i></p> <p>G8. Capacitat de treball en equip</p> <p><i>Sistèmiques</i></p> <p>G14. Capacitat de motivació per la qualitat i per l'assoliment</p> | <p><i>Competències Específiques d'Enginyeria en Informàtica</i></p> <p>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.</p> <p>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.</p> <p><i>Competències Específiques Comunes a la branca de Telecomunicació</i></p> <p>T7. Capacitat de conèixer i utilitzar l'arquitectura i metodologia de disseny, verificació i validació de programari.</p> <p>T8. Capacitat de realitzar programació en temps real, concurrent, distribuïda i basada en esdeveniments, així com el disseny d'interfícies persona-computador.</p> <p><i>Competències de tecnologia específica: Sistemes Audiovisuals</i></p> <p>AU37. Conèixer les tècniques bàsiques de la mineria de dades i textos en la web i capacitat per aplicar-les per a problemes concrets.</p> <p>AU39. Domini de les tècniques avançades de la cerca intel·ligent d'informació en la web.</p> |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

## 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<br>2. Processament del text i la consulta<br>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<br>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<br>2. Jerarquització de pàgines Web<br>3. Escalabilitat | 1. Realització de programes<br>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<br>2. Text estructurat en XML<br>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 4. L'examen de setembre tindrà la mateixa estructura que el de juny.

La nota final de l'assignatura serà la suma d'un 60% de la nota d'examen, un 30% de la nota de pràctiques i un 10% de la nota de seminaris. És imprescindible tenir 4 o més punts a les tres notes parcials per a poder superar l'assignatura, i més de 5 de nota mitjana.

Qualsevol de les notes es pot guardar fins a la convocatòria de setembre. En cap cas d'un curs per a un altre.

## **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, segundaedición. Addison-Wesley, 2011. Pàgina Web: [www.mir2ed.org](http://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](http://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 <sub>1</sub>                                              |                                 |                                              |
| 1. Conceptes de recuperació de la informació | T <sub>1</sub> T <sub>2</sub> T <sub>3</sub> T <sub>4</sub> | P <sub>2</sub>                  |                                              |
| 2. Índexs de cerca en text                   | T <sub>5</sub> T <sub>6</sub>                               | P <sub>3</sub> , P <sub>4</sub> |                                              |
| 3. Cercadors Web                             | T <sub>7</sub> T <sub>8</sub> T <sub>9</sub>                | P <sub>1</sub> , P <sub>5</sub> | S <sub>4</sub>                               |
| 4. Aplicacions                               |                                                             |                                 | S <sub>1</sub> S <sub>2</sub> S <sub>3</sub> |

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 |
| <b>Introducció</b>                                  | 2                         |                           |           |                          |                                            |        |
| <b>1. Conceptes de recuperació de la informació</b> | 6                         | 2                         |           | 4                        | 12                                         |        |
| <b>2. Índexs</b>                                    | 4                         | 4                         |           | 8                        | 8                                          |        |
| <b>3. Cercadors Web</b>                             | 6                         | 4                         | 2         | 8                        | 12                                         |        |
| <b>4. Aplicacions</b>                               |                           |                           | 6         |                          | 9                                          |        |
|                                                     |                           |                           |           |                          |                                            |        |
| <b>Avaluació</b>                                    |                           |                           |           |                          |                                            | 3      |
| <b>Total</b>                                        | 18                        | 10                        | 8         | 20                       | 41                                         | 3      |

**Total: 100**

## 8. Programació d'activitats

- Programació de sessions presencials

|                                       | <b>Dilluns<br/>16:30-18:30</b>                                          | <b>Dimecres<br/>18:30-20:30</b>                                         | <b>Divendres<br/>14:30-16:30</b>                         |
|---------------------------------------|-------------------------------------------------------------------------|-------------------------------------------------------------------------|----------------------------------------------------------|
| <b>Setmana 1</b><br>(9 - 13 abril)    | FESTIU                                                                  | Introducció a la recuperació de la informació (RI)                      | Models de recuperació de la informació                   |
| <b>Setmana 2</b><br>(16 - 20 abril)   | Cercadors en la pràctica: Terrier i Lucene (P101)                       | Cercadors en la pràctica: Terrier i Lucene (P102)                       | Avaluació en RI                                          |
| <b>Setmana 3</b><br>(23 - 27 abril)   | FESTIU                                                                  | Avaluació en RI (P101)                                                  | Processament del text i de la consulta                   |
| <b>Setmana 4</b><br>(30 abril-4 maig) |                                                                         | Avaluació en RI (P102)                                                  | RI en dades XML (S102)<br>Aplicacions (S101)             |
| <b>Setmana 5</b><br>(7 - 11 maig)     | Índexs invertits: construcció                                           | RI en dades XML (S103)<br>Aplicacions (S104)                            | RI en dades XML (S104)<br>Aplicacions (S103)             |
| <b>Setmana 6</b><br>(14 - 18 maig)    | RI en dades XML (S101)<br>Aplicacions (S102)                            | Construcció de índexs invertits (P102)                                  | Construcció de índexs invertits (P101)                   |
| <b>Setmana 7</b><br>(21 - 25 maig)    | Índexs invertits: cerca                                                 | Cerca en índexs invertits (P101)                                        | Cerca en índexs invertits (P102)                         |
| <b>Setmana 8</b><br>(28 maig- 1 juny) | FESTIU                                                                  | Cercadors Web                                                           | Jerarquització de resultats en la Web                    |
| <b>Setmana 9</b><br>(4 - 8 juny)      | Implementació de un Cercador (P101)<br>Jerarquització Web amb AA (S104) | Implementació de un Cercador (P102)<br>Jerarquització Web amb AA (S101) | RI Multimèdia (S101)<br>Jerarquització Web amb AA (S102) |
| <b>Setmana 10</b><br>(11 -15 juny)    | RI Multimèdia (S103)                                                    | RI Multimèdia (S102)                                                    | Escalabilitat en sistemes RI                             |
| <b>Setmana 11</b><br>(18 -22 juny)    | RI Multimèdia (S104)<br>Jerarquització Web amb AA (S103)                | EXÀMENS                                                                 | EXÀMENS                                                  |

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

| <b>Activitat</b>    | <b>Data enunciat</b> | <b>Data lliurament</b> | <b>Data de lliurament de resultats</b> |
|---------------------|----------------------|------------------------|----------------------------------------|
| 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                                   |