# COURSE DATA

## DATA SUBJECT

**Code:** 33847
**Name:** Documentary Management Systems
**Cycle:** Undergraduate Studies
**ECTS Credits:** 6
**Academic year:** 2025-26

## STUDY (S)

| Degree | Center | Acad. year | Period |
|---|---|---|---|
| 1007 - Degree in Information and Documentation | Facultat de Geografia i Història | 3 | Second quarter |

## SUBJECT-MATTER

| Degree | Subject-matter | Character |
|---|---|---|
| 1007 - Degree in Information and Documentation | Information technologies and digital editing | COMPULSORY |

## COORDINATION

CIRILO GIMENO RAMON VICENTE

# SUMMARY

This course covers the functional and technological foundations of computer systems designed to store documents in digital format, specially text documents, and the base technologies that can be found in the implementation of this type of applications (Document Management Systems – DMS).

With this view, students will be presented the information elements fundamentals in the scope of document management, DMS processes for indexing and querying documents, the general functional architecture of a DMS, several information retrieval models, some algorithms for pre-processing, indexing, querying and matching texts, and assessing the overall performance of any DMS and queries processes and results.

# PREVIOUS KNOWLEDGE

## RELATIONSHIP TO OTHER SUBJECTS OF THE SAME DEGREE

There are no specified enrollment restrictions with other subjects of the curriculum.

## OTHER REQUIREMENTS

It is recommended to have taken Informática 1, Informática 2, Sistemas de Representación de Información y Conocimiento, y Arquitectura de Información en la Web prior to taking this course.

## COMPETENCES / LEARNING OUTCOMES

### 1007 - Degree in Information and Documentation

Be able to apply critical reasoning to the analysis and assessment of alternatives.

Be able to design information products and services in any field and by any means of dissemination (electronic edition) according to the information and training needs detected in a community of users.

Be able to learn independently.

Be able to search and retrieve information by methods that meet the expectations and needs of users in optimal conditions of cost and time.

Be able to use and put into practice methods, techniques and computer tools (hardware or software) for the design, implementation, development and operation of information systems.

Be able to work in a team and to integrate into multidisciplinary teams.

Be sensitive to environmental issues, sustainability and human rights.

Capacity to write analytical reports and summaries with regard to management and organisation of information.

Demonstrate organisational and planning skills.

Have computer skills related to the field of study.

Have decision-making capacity.

Have problem-solving skills.

Have skills for information management.

Know, use and apply information and communication technologies applied to the storage, use, management, handling, distribution and exploitation of data, information and knowledge.

Know, use and apply the computer and telecommunications tools that support the development of the set of skills that must be acquired in the training process.

Know a foreign language.

Show creativity.

Show ethical commitment in the relationships with users and in information handling.

Show management and leadership skills.

Show motivation for quality.

Understand, design and apply models for data and information representation, and mechanisms for data extraction and exploitation and for information retrieval.

## DESCRIPTION OF CONTENTS

### 1. Introduction and basic concepts

Lecture 1. Introduction

- Information needs
- Concept of information and document
- Information representation
- Concept of Information Retrieval. Functional scheme.

### 2. Information Retrieval Models

Lecture 2. Information Retrieval Models

- The IR ideal model
  - Definition of precision and recall
  - Behaviour of the ideal system
- Classical IR models
  - Boolean model
  - Vector-space model
  - Probabilistic mocel
- More models

Lecture 3. Functional architecture of a DMS

- Functional description
  - Document indexing
  - Full search
  - Selective dissemination of information
  - Feedback
- DMS structure
- Query capabilities
  - Types of queries (terms, phrases, proximity, fuzzy, wildcards)
  - Hits management (ranking, zoning, highlighting)

- General (vocabulary navigation, query storage, thesaurus management)

## 3. Processes, data strctures and algorithms

Lecture 4. Document management and retrieval

- Document indexing
  - Indexes implementation
- Query operations
  - Boolean queries resolution
  - Vector-space simmilarity calculation
  - Sequential search
- Feedback implementation

Lecture 5. Text processing

- Lexical analysis
- Stopwords
- Stemming
- Thesaurus and clustering

## 4. Evaluation of DMS

Lecture 6. Evaluation of DMS

- Systems evaluation
  - Systems and components to be evaluated
- Query evaluation
  - Evaluation measurements
- The TREC conference as an example of evaluation

# WORKLOAD

## PRESENCIAL ACTIVITIES

| Activity | Hours |
|---|---|
| Theory | 37,50 |
| Laboratory | 22,50 |
| **Total hours** | **60,00** |

## NON PRESENCIAL ACTIVITIES

| Activity | Hours |
|---|---|
| Attendance at other activities | 2,00 |
| Individual or group project | 23,00 |
| Independent study and work | 25,00 |
| Preparation of lessons | 25,00 |
| Preparation for assessment activities | 15,00 |
| Resolution of case studies | 0,00 |
| **Total hours** | **90,00** |

## TEACHING METHODOLOGY

LECTURES:

The lectures will be based on active lectures where every 20/25 minutes will be introduced in any activity that requires the involvement of students, so that 1) they can do an activity based on the content they have just learnt, 2) they recover the level of attention to the next block.

LECTURES PREPARATION:

Students will have to prepare the lecture content, following the plan of the course. To do this they will use the literature suggested by the teacher as well as materials provided or/and any other directions given by the lecturer.

PREPARATION OF PRACTICAL WORK:

To better assimilate the contents of the lectures, practical sessions will be conducted in the laboratories. Attendance at practical sessions is mandatory and will be verified by the lecturer in charge of the session. Those students that (for a justified reason) cannot attend the practical sessions should contact the lecturer before the beginning of the first session. The results of these activities must be submitted to the lecturer in charge of the group during the course and in the terms established by the teacher. Students are expected to do/prepare some of these activities at home. Attendance to the practical sessions is mandatory.

TEAM WORK:

A set of problems will be proposed that should be solved in teams of 3 to 6 persons. Each member of the group will be graded both the joint mark of the group as the individual mark of each member.

TUTORIALS:

a) Scheduled tutorials:

Students will work in small groups on some of the more complex concepts introduced in the lectures. They will be provided with a range of activities / problems that will be solved with the help of the lecturer

b) Unscheduled tutorials:

There will be some tutoring hours per week where the lecturer will assist students to clarify concepts or doubts that have arisen during the lectures.

COMPLEMENTARY ACTIVITIES:

There will be a seminar which will address in detail some of the topics discussed throughout the course. The seminar duration will be 2 hours (in-person).

## EVALUATION

The evaluation of the subject will follow a scheme of continuous assessment. The following aspects will be considered:

1. Written tests: there will be only one final written test of theoretical - practical character. The minimum score a student must achieve to pass the course is 5 out of 10. The grade obtained in this test represent 50% of the final grade.
2. Preparation of lectures, activities and problems that arise in the context of the lectures will have 15% of the final grade. All activities are mandatory for completion of the course evaluation and are not recoverable in the second call.
3. Practical work: The grade in this section represent 30% of the final grade. All practical work is required for completion of the course evaluation and is not recoverable in the second call.
4. Team work: In the evaluation of team work will be graded both the joint mark of the group as the individual mark of each member. The grade in this section represent 5% of the final grade. All supervised work in a team is required to conduct the evaluation of the subject and is not recoverable in the second call.
5. The composition of the final grade will follow, in summary, the following table:

- Exam: 50%
- Preparation of lectures: 15 %
- Practical work: 30 %
- Teamwork: 5 %

Students who do not submit any of the required individual, practical, or team assignments throughout the course will be marked as absent in the first examination period, regardless of the grades they may obtain in the remaining activities. They will have the opportunity to participate in the second examination period with the grades obtained in the activities they have submitted, without the possibility of recovery.

This assessment starts from the premise that teaching at the University of Valencia is, by definition, on-campus lecture delivery method. In this sense, the student should be aware that attendance at both the theoretical and practical lectures is essential for proper monitoring of the contents of the course. The student must also consider the possibility to enroll part time when it is unable to attend all courses (60

credits). However, there is an exception for those students that justify it and request it. They have the possibility of being assessed without attending to all or part of the lectures. For these cases, students should proceed as follows:

- At the beginning of the course, student should inform to lecturer responsible for the course, the incidence that makes her/him unable to attend the class. This must be adequately justified in documentary form.
- The lectures in charge, in the light of this information, will decide the possibility of evaluation without full or partial assistance to the lectures.

Students who are in this situation must submit for evaluation all work required by the lecturer (not necessarily the same to those required for the course) and may also be called to defend them orally to the lecturer, and conduct a knowledge test. The weight of the final grade work will be 50% and the test the remaining 50% knowledge.

## REFERENCES

- Modern Information Retrieval. R. Baeza-Yates, B. Ribeiro-Neto. Addison-Wesley. 1999.
- Introduction to Information Retrieval. C.D. Manning, P. Raghavan, H. Schütze. 2008. http://nlp.stanford.edu/IR-book/information-retrieval-book.html
- Information Retrieval. Data structures and algorithms. W.B. Frakes, R. Baeza-Yates. Prentice-Hall. 1992
- Information Retrieval. C. J. van Rijsbergen. 1979. http://www.dcs.gla.ac.uk/Keith/Preface.html
- Bases de Datos Documentales. E. Abadal, L. Codina. 2005.
- "Text Information Retrieval Systems" C.T. Meadow. Academic Press. 1992.
- "Automated Information Retrieval. Theory and Methods" V. Frants, J. Shapiro, V. Voiskunskii. Academic Press. 1997.
- "Information Retrieval Systems. Theory and Implementation" G. Kowalsky. Kluwer Academic Publishers. 1997.
- Information Storage and Retrieval. R. Korfhage. Wiley. 1997.

- Information Retreival Resources: http://nlp.stanford.edu/IR-book/information-retrieval.html