

## INFORMATION RETRIEVAL AND ACCESS

Academic Year: ( 2016 / 2017 )

Review date: 24-03-2015

Department assigned to the subject: Department of Computer Science and Engineering

Coordinating teacher:

Type: Electives ECTS Credits : 6.0

Year : Semester :

## STUDENTS ARE EXPECTED TO HAVE COMPLETED

As an optional complement, it is recommended to have taken the course Information Engineering, since said subject complements some of the concepts dealt with during the class.

## COMPETENCES AND SKILLS THAT WILL BE ACQUIRED AND LEARNING RESULTS.

1. Acquisition of the relevant knowledge of the modern techniques of Information Retrieval. The approach is mainly focused in the Web and the Intranets environments.
2. Knowledge of the methodologies used in current systems (Natural Language Processing, Data Mining and machine learning, Knowledge Discovery, Information extraction, Knowledge Organization Systems, ¿).
3. Use of existing applications for information retrieval

## DESCRIPTION OF CONTENTS: PROGRAMME

Information Retrieval and Access

- Retrieval Algorithms
- Assigned Attribute Weights
- Retrieval Models

Internet Retrieval

- Types of search engines
- Positioning Algorithms
- The future of Information Retrieval: Ontologies and Web Semantics

Quality Measurement in Document Retrieval

- Types of measurements
- Problems
- Test collections

Corpora

- Text Corpus
- Corpus as Linguistic Resources

Automatic Indexing

- Types of Indexing
- Classic methodology: Tokens and Filtering

NLP based analysis: Analyzers, Resources and Problems

Information Retrieval Languages

Improving retrieval systems:

- Name entity recognition and information extraction
- Text Mining
- Knowledge Organization Systems

## LEARNING ACTIVITIES AND METHODOLOGY

- I. Web Page Development with Search Engine Optimization criteria
- II. Developing a sistem aimed to Plagiarism Detection
- III. Design and implementation of a Search engine system:
  - a) Search Engine implemented with NLP and Information Science techniques
  - b) Improving the search engine with entity recognition and text mining capabilities
  - c) Improving the search engine with semantic and lexical networks
  - d) Evaluation by test collections

## ASSESSMENT SYSTEM

Acquisition of theoretical and practical knowledge of the course topics. Quality and relevant characteristics of the practical notebook and the oral presentations done during the course (60% final score). Final test regarding knowledge acquired during the classes (40%).

