Advanced Techniques For Information Retrieval

Academic Year: (2019 / 2020)   **Review date: 27-04-2020**

Department assigned to the subject: Computer Science and Engineering Department

Coordinating teacher: MORATO LARA, JORGE LUIS

Type: Electives ECTS Credits : 3.0

Year : 1 Semester : 0

OBJECTIVES

BASIC COMPETENCIES
CB9   Students should know how to communicate their conclusions, knowledge and reasoning to specialized and non-specialized audiences, in a clear and unambiguous way

GENERAL COMPETENCIES
CG5   Acknowledge the growing importance of team-working in the labor market and show adaptability and integration capabilities in different work environments, keeping relationships and fluid communications.

CG8   To value the quality of the work undertaken when planning, organizing and developing the activities related with the grade. The student should be able to show initiative, creativity and sense of the responsibility, keeping the interest during the process and showing commitment for results obtained.
CG9   Integrate knowledge, report informed judgments and communicate conclusions, grounded in specialized knowledge and reasoning, to specialized and non-specialized audiences, in a clear and unambiguous way
CG11 Ability to interpret, apply and innovate methodologies, technologies, policies, analysis, and information management and retrieval methods

SPECIFIC COMPETENCIES
CE1   Learn and analyze the current state and the future perspectives and application of these technologies in libraries and archives.
CE6   Using metadata vocabularies and other semantic schema models for managing digital documents.
CE7   Data visualization, using temporal techniques, geospatial, thematic and network analysis.

LEARNING RESULTS
In this course is specially important all the techniques that facilitates information exchange and publication of digital documents, specially in regard with Semantic Web technologies.
After completing the course the student should know:
1. Evaluate the main Information Retrieval Systems, emphasizing retrieval on the Web and the Semantic Web.
2. Know how-to clean and augment data in BigData lyfecycle
3. Evaluate retrieval systems
4. Know how-to disseminate of documentation on the Web: search engine positioning and management
5. Learn formalisms and strategies to improve the interoperability and organization of documents
6. Know how to use the main retrieval languages in databases and the Semantic Web.

DESCRIPTION OF CONTENTS: PROGRAMME

Common contents in regard with the other subjects:
- Using information retrieval systems and semantic schemas.

Specific contents to the subject:
- Search in different web types: Semantic web, social web, data web, question-answering web, deep web, dark web, commercial web.
- Information retrieval systems for Big Data, documents (scrapers) and Linked Data
- Evaluation metrics for information retrieval .
- Knowledge management model of knowledge representation and organization of information and

semantic Interoperability
- Techniques of positioning and SEO tools that affect the indexing and retrieval
- Fundamentals in information retrieval languages: SQL, SPARQL, Xpath, Regex

Assigments:
- Academic activities based on exercises and problems.
- Seminar/Workshop based on use cases

LEARNING ACTIVITIES AND METHODOLOGY

LEARNING ACTIVITIES IN THE STUDY PLAN
AF1   Individual work related with theoretical contents and practicalities delivered by the professor.
AF2   Individual work for problem resolution and case study
AF3   Theoretical and practical face-to-face classes
AF4   Tutorial class
AF5   Teamwork
AF6   Active involvement in forums in the platform
AF7   Self-assessment tests

Activity cod	Total hours		Classroom classes	%	
AF1	125	(32)	0		0
AF2	80	(30)	0		0
AF3	12	(3)	12 (3)		100 (100)
AF4	10	(2)	0		0
AF5	124	(18)	0		0
AF6	5	(2)	0		0
AF7	4	(3)	0		0
TOTAL	360	(90)	12 (3)	3,3 (3,3)	

TEACHING METHODOLOGY
MD1 Class presentations by the teacher, with computer and audiovisual resources, in these classes the basic concepts in the subject will be explained and a basic bibliography will be provided.
MD2. Critical review of the text suggested by the teacher: articles, reports, manuals and research papers, to complement the teacher's material.
MD3. Resolution of practical case studies and problems, to be solve in teams or individually.
MD5. Report writing individually or in teams.
MD6. Reading teaching theoretical and practical materials

TUTORIAL CLASES
Tutorial classes will be scheduled according the regulations provided by the University. They will be published in the e-leaning platform (Aula Global). There will be two types of tutorial classes, face-to-face and online. Students will ask for individual tutorial classes in different hours to those published.

ASSESSMENT SYSTEM
SE1   Class participation and forums participation
SE2   Individual or team works during the course
SE3   Evaluation questionnaries
SE4   End-of-term examination
SE5   Presentation and quality of the Master Dissertation Thesis

* End-of-term examination will be done in the presence-based modality, in the university Carlos III or related center. In order to pass the subject the student should pass this exam.

Assessment System

SE1 Forum and class participation 5
SE2
Individual work SEO positioning and search engines: 10
Individual work Evaluation metrics: 10
Individual work SQL: 10
SE3
- Self-assessment tests: 5
- SPARQL Questionary: 10
SE4
- Final exam: 20
- Team work: 30

A minimum mark of 5 in the final exam is needed in order to pass the subject.
In case of failing in the continual assessment process, the maximum calification in the final exam will a 60%, the extraordinary exam will have a weight of 75%

% end-of-term-examination: 20

% of continuous assessment (assigments, laboratory, practicals…):   80

BASIC BIBLIOGRAPHY

- Baeza-Yates, Ricardo Modern Information Retrieval, ACM Press, 2011

- Dean Allemang, James Hendler Semantic Web for the Working Ontologists: Effective Modelin in RDFS and OWL, Elservier, 2011
- Verborgh, Ruben, De Wilde, Max Using OpenRefine: the essential OpenRefine guide that takes you from data analysis and error fixing to linking your dataset to the Web, Packt Publishing, 2013

BASIC ELECTRONIC RESOURCES

- Cody Burleson . SPARQL Query Examples: https://wiki.base22.com/display/btg/SPARQL+Query+Examples

- Enipedia . OpenRefine Tutorial: http://enipedia.tudelft.nl/wiki/OpenRefine_Tutorial

- W3schools . SQL Tutorial: https://www.w3schools.com/sql/