Advanced Techniques For Information Retrieval

Academic Year: (2023 / 2024) **Review date: 19/05/2023 19:24:34**

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 : 2

OBJECTIVES

COMPETENCIES
C1 Be able to undertake projects that develop lines of technological innovation in libraries and archives and documentation centers.
C2 Perform technological advisory and consulting tasks for the implementation of basic services of creation, digitization, digital preservation, treatment, dissemination and preservation of information in digital format, treatment, dissemination and preservation of information in digital format.
C3 Integrate future perspectives in the activity developed in libraries and archives, in terms of technological advances and updating of knowledge.
knowledge.
C5 Use metadata vocabularies and other semantic schema models for the treatment of digital documents.

DESCRIPTION OF CONTENTS: PROGRAMME

Topics common to the subject:
- Use of information retrieval systems and semantic schemas.

Specific topics of the subject:
Theory
- Information retrieval systems applied to big data, documents and linked data.
- Evaluation metrics for information retrieval.
- Knowledge management through knowledge representation models and information organization and semantic interoperability.
- Positioning techniques and SEO tools that affect indexing and retrieval.
- Main information retrieval languages

Practice
- Practice based on exercises and problems.
- Seminars-workshops with case studies.

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

% end-of-term-examination/test:	20
% of continuous assessment (assigments, laboratory, practicals…):	80

SE1 Participation in class and forums in the virtual educational platform.
SE2 Individual or group work carried out during the course.
SE3 Completion of evaluable and graded questionnaires
SE4 Final exam or paper

* The final exam will take place at Universidad Carlos III de Madrid, and must be passed in order to pass the course.

EVALUATION SYSTEM

SE1 Participation in forums and face-to-face sessions 5
SE2 45
Individual practice Positioning
Individual practice Evaluation metrics
Individual practice SQL

SE3 15
Performance of evaluation tests
SE4
- Final on-site exam using the Aula Global platform: 20
- Completion and Delivery to (within the time frame reserved for the final on-site test): 15
(if applicable, data acquisition with scrapers and SPARQL and cleaning with Refine): 15)

A five in the final exam and a five in the continuous evaluation is the minimum grade to pass the course.

In the extra exam session the maximum grade will be 75%.

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

ADDITIONAL BIBLIOGRAPHY

- Anne Ahola Ward The SEO battlefield : winning strategies for search marketing programs, OReilly, 2017

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

- Google . Search Engine Optimization (SEO) Starter Guide: https://support.google.com/webmasters/answer/7451184?hl =en

- MOZ . The beginner's guide to SEO: https://moz.com/beginners guide to SEO

- Manning, Christopher D; Raghavan , Prabhakar ; Schütze , Hinrich . Introduction to Information Retrieval: https://nlp.stanford.edu/IR book/information retrieval book.html

- Stardog . Tutorial Sparql: https://www.stardog.com/tutorials/

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