Course Fact sheet

Master in Informatics and Computing Engineering
Markup Languages and Document Processing
Instance: 2016/2017
—

Institutional page

General information

Course Unit: Markup Languages and Document Processing
Code: EIC0107
Programmes: MIEIC, 4º
Academic Year: 2016/2017
Semester: 2S
Credits: 6
ECTS: 6
Hours/Weeks: 3 TP
Teachers: João Correia Lopes

Teaching language

English.

Objectives

BACKGROUND

The "Markup Languages and Document Processing" unit assumes as its context the widespread use of markup languages to represent semi-structured information and the existence of standardized tools for their treatment.

OBJECTIVES

1. Make the students aware of the multiple non-trivial applications of markup languages;
2. Make the students familiar with the technologies of processing and storage of semi-structured information;
3. Explore the technologies in defining an annotation language and in the development of an application.

Skills and learning outcomes

On completion of this course, the student should be able to:

- Identify the use of markup languages in documents, in data repositories and in applications;
- Create models for XML documents;
- Tell the difference between a standardised language for an application domain and the ad-hoc uses of markup languages;
- Design a markup language to support a document type or the data for an application;

- Evaluate and compare XML-based and other solutions to support application data interchange;
- Design XSL stylesheets to transform documents;
- Interpret the results of document processing with XML-based technologies;
- Use a native XML database and take advantage of its query functions;
- Compare data organisation in markup models with the relational model and translate data between both models;
- Generate a markup model for data in an application domain, store a dataset and query it;
- Compare markup languages to other document and data representations with respect to the support to data preservation along technological change;
- Relate web documents with the metadata that describes or links them;
- Compare semantic web based services with simpler approaches to resource description;
- Develop an XML-based prototype application involving the use of a dialect and document processing.

Work mode

Classroom.

Previous knowledge

Students must have basic skills in markup languages and technologies.

Program

- Design of markup languages. Analysis of existing languages for different domains.
- Querying XML documents. The XPath language. The XQuery language.
- XML document transformation with XSLT. XML presentation with XSL-FO.
- XML processing in applications. The standard interfaces DOM and SAX.
- Storage of XML in native XML databases. XML storage in object-relational databases.
- Introduction to the Semantic Web. RDF, OWL, Linked Open Data.

Main bibliography

- Anders Møller, Michael I. Schwartzbach; *An Introduction to XML and Web Technologies*, Addison Wesley Professional, 2006. ISBN: 0321269667 Biblioteca

Complementary bibliography

- Neil Bradley, The XML Companion, Addison-Wesley, 3rd Edition, 2002, ISBN: 0-201-77059-8. Biblioteca
- José Carlos Ramalho, Pedro Henriques; *XML e XSL — da Teoria à Prática*, FCA Editora, 2002, ISBN: 972-722-347-8. Biblioteca
- Elliotte Rusty Harold, W. Scott Means, *XML in a Nutshell*, O'Reilly, Third Edition, 2004, ISBN: 0-596-00292-0. Biblioteca
- Grigoris Antoniou and Frank van Harmelen, *A Semantic Web Primer*, 2nd edition, The MIT Press, 2008, ISBN: 9780262018289

Teaching procedures and learning activities

Lectures include theoretical presentation of the course subjects and practical sessions where proposed research topics are discussed with the students and practical coursework is reported.

Support software

- eXist, MongoDB
- oXygen

Keywords

Physical sciences > Computer science > Informatics

Evaluation type

Distributed evaluation without final exam

Evaluation and occupation components

The practical work is broken down into three components with the delivery of a small scientific paper and a presentation and discussion in a workshop.

The theoretical concepts are evaluated through the individual response to 4 questionnaires.

Evaluation components

Description	Type	Time (hours)	Conclusion date
Teaching classes (estimated)	lectures	42	—
Quiz #1: storage (Q1)	Questionnaire	0	01/03/2017
Project proposal (D1)	Work	27	08/03/2017
Quiz #2: XQuery (Q2)	Questionnaire	0	22/03/2017
Quiz #3: XSLT (Q3)	Questionnaire	0	05/04/2017
Architecture and prototype (D2)	Work	33	19/04/2017
Quiz #4: RDF (Q4)	Questionnaire	0	17/05/2017
XML application (D3)	Work	60	24/05/2017
	Total:	162	—

Occupation components

Practical work 120
Classes 42

Admission to exams

n/a

Final grade

Classification = 70% Practical assignments + 30% Questionnaires, where: Practical assignments = 20% D1 + 30% D2 + 50% D3, Questionnaires = (Q1 + Q2 + Q3 + Q4) / 4

The practical work will be assessed through the documentation submitted up to their due dates established in the course plan, the participation in the presentations and through the developed XML-based application.

The individual questionnaires last for 20 minutes and take place during class.

The minimum required to pass the course is 40% in each of the deliverables and presentations of practical assessment (D1 to D3) and 40% on average of the Q1 to Q4 questionnaires related theoretical subjects. There is no recovery examination.

This course, given its nature, is not eligible for evaluation in a single moment, so the practical work evaluation cannot be replaced by taking an exam.

Special assignments

There are no special works or tests.

Special evaluation (TE, DA, ...)

Students taking exams under special regimes are expected to previously submit the assignments required for this course at the same time as regular students.

Improvement of final/distributed classification

Students may improve the mark in the course's next edition.

– JCL