

Caracterização da Unidade Curricular / Characterization of the Curricular Unit

Designação da Unidade Curricular (UC) / Title of Curricular Unit (CU): Processamento de Informação / Information Processing

Área científica da UC / CU Scientific Area: Informática / Computer Science

Semestre / Semester: 3º

Número de créditos ECTS / Number of ECTS credits: 6

Carga horária por tipologia de horas / Workload by type of hours: TP: 22,5; PL: 22,5; OT: 6; O: 9

Carga letiva semanal / Weekly letive charge: 3h

Objetivos de aprendizagem (conhecimentos, aptidões e competências a desenvolver pelos estudantes):

No final da unidade curricular os alunos deverão ser capazes de propor e desenvolver soluções para representação de informação, extração de informação a partir de conjuntos de dados (datasets) e de utilizar técnicas e ferramentas para identificação de padrões em grandes volumes de dados.

Intended learning outcomes (knowledge, skills and competences to be developed by the students):

At the end of the curricular unit, students should be able to propose and develop solutions for information representation, extract information from datasets (datasets) and use techniques and tools to identify patterns in large volumes of data.

Conteúdos programáticos:

1. Conceitos base do Processamento de Informação

- 1.1. Registo
- 1.2. Codificação
- 1.3. Armazenamento
- 1.4. Recuperação

2. Linguagem natural vs. Linguagens estruturadas

- 2.1. Formas de representação de informação
- 2.2. Ambiguidade da linguagem natural
- 2.3. Noções de semântica
- 2.4. Mapas mentais
 - 2.4.1. Conceitos e Concetualização
 - 2.4.2. Ferramentas de Representação
 - 2.4.3. Noção de Ontologia

3. Documentos e anotação

- 3.1. O papel da meta-informação
- 3.2. Marcas
- 3.3. Códigos de Carateres
- 3.4. Classificação

4. Linguagens de marcação e interpretadores

- 4.1. HTML, XML como linguagens de marcação
- 4.2. Organização e formatação dos documentos
- 4.3. Marcas, atributos, identificadores, validação

5. O papel das linguagens de scripting

- 5.1. Expressões regulares
- 5.2. Extração de texto natural
- 5.3. Exploração e extração de informação de datasets

6. Aplicações e algoritmos para processamento de informação

Syllabus:

1. Basic concepts of Information Processing

- 1.1. Registration
- 1.2. Coding
- 1.3. Storage
- 1.4. Recovery

2. Natural language vs. structured languages

- 2.1. Forms of information representation
- 2.2. Ambiguity of natural language
- 2.3. Notions of semantic
- 2.4. Mind maps
 - 2.4.1. Concepts and Conceptualization
 - 2.4.2. Representation Tools
 - 2.4.3. Notion of Ontology

3. Documents and annotation

- 3.1. The role of meta-information
- 3.2. Tags

3.3. Character codes

3.4. Classification

4. Markup languages, Parsers, Interpreters

4.1. HTML, XML as markup languages

4.2. Organization and formatting of documents

4.3. Tags, Attributes, identifiers, validation

5. The role of scripting languages

5.1. Regular expressions

5.2. Processing of natural language

5.3. Exploration and extraction of information from datasets

6. Applications and algorithms for information processing

**Sem Validade
Administrativa**