COURSE OUTLINE

1. GENERAL

<table>
<thead>
<tr>
<th>SCHOOL</th>
<th>SCHOOL OF ENGINEERING</th>
</tr>
</thead>
<tbody>
<tr>
<td>DEPARTMENT</td>
<td>DEPARTMENT OF COMPUTER ENGINEERING AND INFORMATICS</td>
</tr>
<tr>
<td>LEVEL OF COURSE</td>
<td>UNDERGRADUATE, CORE ELECTIVE</td>
</tr>
<tr>
<td>COURSE CODE</td>
<td>CEID_NE5597</td>
</tr>
<tr>
<td>SEMESTER OF STUDIES</td>
<td>WINTER</td>
</tr>
<tr>
<td>COURSE TITLE</td>
<td>INFORMATION RETRIEVAL</td>
</tr>
</tbody>
</table>

INDEPENDENT TEACHING ACTIVITIES

if credits are awarded for separate components of the course, e.g. lectures, laboratory exercises, etc. If the credits are awarded for the whole of the course, give the weekly teaching hours and the total credits

<table>
<thead>
<tr>
<th></th>
<th>TEACHING HOURS PER WEEK</th>
<th>ECTS CREDITS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lectures</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Labor. Ex.</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Recitation sections</td>
<td>2</td>
<td>2</td>
</tr>
</tbody>
</table>

Add rows if necessary. The organisation of teaching and the teaching methods used are described in detail at (d).

COURSE TYPE

skill development

PREREQUISITE COURSES:
There are no prerequisites. Recommended prerequisite knowledge are the courses of data structures (CEID_NY233), algorithms (CEID_NY205) and data base development (CEID_NY344)

TEACHING AND ASSESSMENT LANGUAGE:
Greek. Instruction may be given in English if foreign students attend the course.

THE COURSE IS OFFERED TO ERASMUS STUDENTS:
YES

COURSE WEBPAGE (URL):
https://eclass.upatras.gr/courses/CEID1037/

2. LEARNING OUTCOMES

Learning outcomes
The course learning outcomes, specific knowledge, skills and competences of an appropriate level, which the students will acquire with the successful completion of the course are described.
Consult Appendix A
- Description of the level of learning outcomes for each qualifications cycle, according to the Qualifications Framework of the European Higher Education Area
- Descriptors for Levels 6, 7 & 8 of the European Qualifications Framework for Lifelong Learning and Appendix B
- Guidelines for writing Learning Outcomes

At the end of this course, the student should be able to:
1. explain basic information storage and retrieval concepts and understand the issues that are specific to the design of efficient information retrieval systems.
2. explain the differences between alternative information retrieval models, and analyze why a particular model is appropriate for a specific application.
3. understand the basic design issues that are involved in the development of an efficient web search engine.
4. design and implement a small to medium size information storage and retrieval system.

At the end of the course the student will have further developed the following skills/competences:
1. ability to understand the various concepts of information retrieval systems, and how these are interrelated with the performance to a working system
2. ability to apply methodologically these concepts in order to design and implement efficient working systems ability to work cooperatively in order to solve problems that arise during the construction of a
full-fledged working software system.

1. General Abilities

Taking into consideration the general competences that the degree-holder must acquire (as these appear in the Diploma Supplement and appear below), at which of the following does the course aim?

- Search for, analysis and synthesis of data and information, with the use of the necessary technology
- Decision-making
- Working independently
- Teamwork
- Working in an international environment
- Production of new research ideas

2. COURSE CONTENT

1. Introductory notions (user modeling, document logical representation, retrieval process).
2. Performance evaluation metrics (recall, precision, average precision, R-precision, precision histograms, NDCG metric, harmonic median, user oriented metrics).
3. Information retrieval modeling.
4. Set-oriented models (boolean models, fuzzy set model, extended boolean model), algebraic models (vector space models, latent semantic indexing model, topic models), probabilistic models (classical and language models).
5. Web information retrieval and its peculiarities.
7. Machine Learning Techniques in Information Retrieval (Learning to Rank, Vector representation of words and Large Language Models (word embeddings such as word2vec, CBOW, skipgram, LSTM, Transformers, BERT, GPT)), Search Engines against Reasoning Engines.
8. Storage Techniques in Distributed Information Retrieval (MapReduce, Apache Spark)
9. Indexing structures (inverted files, signature files, bitmaps).
10. Full indexing structures in main memory (suffix trees, suffix arrays, acyclic directed graphs (DAWG) for strings), and in secondary memory (supra-suffix array, prefix B-tree, string B-tree).
11. Compression algorithms for text and for indexing structures.
12. Text Mining

3. TEACHING AND LEARNING METHODS - ASSESSMENT

<table>
<thead>
<tr>
<th>TEACHING METHOD</th>
<th>Face-to-face, Distance learning, etc.</th>
</tr>
</thead>
<tbody>
<tr>
<td>USE OF INFORMATION AND COMMUNICATION TECHNOLOGIES</td>
<td>We use Information and Communications Technology in communication with students. We use e_class, e_mail.</td>
</tr>
<tr>
<td>TEACHING ORGANIZATION</td>
<td>The manner and methods of teaching are described in detail. Lectures, seminars, laboratory practice, fieldwork, study and analysis of bibliography, tutorials, placements, clinical practice, art workshop, interactive teaching, educational visits, project, essay writing, artistic creativity, etc.</td>
</tr>
<tr>
<td>Дрαστηριότητα</td>
<td>Φόρτος Εργασίας Εξεμήνου</td>
</tr>
<tr>
<td>Lectures</td>
<td>50</td>
</tr>
<tr>
<td>Recitation sections</td>
<td>50</td>
</tr>
<tr>
<td>Laboratory exercises</td>
<td>50</td>
</tr>
<tr>
<td>Total number of hours for the Course</td>
<td>150</td>
</tr>
</tbody>
</table>
The student’s study hours for each learning activity are given as well as the hours of non-directed study according to the principles of the ECTS.

**STUDENT ASSESSMENT**

Description of the evaluation procedure

Language of evaluation, methods of evaluation, summative or conclusive, multiple choice questionnaires, short-answer questions, open-ended questions, problem solving, written work, essay/report, oral examination, public presentation, laboratory work, clinical examination of patient, art interpretation, other

Specifically-defined evaluation criteria are given, and if and where they are accessible to students.

- composing and presenting an essay by groups of 1-2 students, concerning the critical presentation, implementation and analysis of algorithms in a set of scientific papers that deals with a specific topic of the course (45% of the final mark, taken into account only when the student secures a minimum mark in the final written examination),

- written examination (55% of the final mark).

5. **RECOMMENDED LITERATURE**