

COURSE OUTLINE

(1) GENERAL

SCHOOL	Science and Technology		
ACADEMIC UNIT	Science and Technology		
PROGRAMME OF STUDIES	MSc in Data Science		
LEVEL OF STUDIES	Postgraduate		
COURSE CODE	DSC06	SEMESTER	2
COURSE TITLE	Information Retrieval		
COURSE TYPE <i>Elective, compulsory</i>	Elective		
INSTRUCTOR(S)	Prof. Panayiotis Bozanis		
INDEPENDENT TEACHING ACTIVITIES <i>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</i>	WEEKLY TEACHING HOURS	CREDITS	
	3	6	
<i>Add rows if necessary. The organisation of teaching and the teaching methods used are described in detail at (d).</i>			
TEACHING ACTIVITIES BREAKDOWN	WEEKLY HOURS		
Theory	2,3		
Recitation	0,7		
Lab	0		
<i>Add rows if necessary. The organisation of teaching and the teaching methods used are described in detail at (d).</i>			
COURSE TYPE <i>general background, special background, specialised general knowledge, skills development</i>	Specialized Knowledge		
PREREQUISITE COURSES:	-		
LANGUAGE OF INSTRUCTION and EXAMINATIONS:	English		
IS THE COURSE OFFERED TO ERASMUS STUDENTS	Yes		
COURSE WEBSITE (URL)	https://elearn-ucips.ihu.gr/		

(2) LEARNING OUTCOMES

<p>Learning outcomes</p> <p><i>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.</i></p> <p><i>Consult Appendix A</i></p> <ul style="list-style-type: none"> • <i>Description of the level of learning outcomes for each qualifications cycle, according to the Qualifications Framework of the European Higher Education Area</i> • <i>Descriptors for Levels 6, 7 & 8 of the European Qualifications Framework for Lifelong Learning and Appendix B</i> • <i>Guidelines for writing Learning Outcomes</i>
<p>On completing the course, the student will be able to:</p> <ul style="list-style-type: none"> • Understand key concepts of information retrieval techniques.

- Apply information retrieval principles to locate relevant information in large collections of documents.
- Apply information retrieval models.
- Understand and deploy efficient techniques for the indexing of documents.
- Analyse the performance of retrieval systems.
- Know the basic concepts of modeling and searching the Web.

General Competences

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?

<i>Search for, analysis and synthesis of data and information, with the use of the necessary technology</i>	<i>Project planning and management</i>
<i>Adapting to new situations</i>	<i>Respect for difference and multiculturalism</i>
<i>Decision-making</i>	<i>Respect for the natural environment</i>
<i>Working independently</i>	<i>Showing social, professional and ethical responsibility and sensitivity to gender issues</i>
<i>Team work</i>	<i>Criticism and self-criticism</i>
<i>Working in an international environment</i>	<i>Production of free, creative and inductive thinking</i>
<i>Working in an interdisciplinary environment</i>
<i>Production of new research ideas</i>	<i>Others...</i>

- Search for, analysis and synthesis of data and information, with the use of the necessary technology
- Decision Making
- Production of free, creative, and inductive thinking

(3) SYLLABUS

The course overviews basic principles and techniques of information retrieval. The students will familiarize themselves with fundamental data models, organization schemes and data processing - searching techniques that allow for efficient information retrieval over a collection of documents located in the world wide web. The topics covered include:

- Boolean Model.
- Term Vocabulary.
- Indexing.
- Vector Space Model.
- Scoring.
- IR Evaluation.
- Web Search and Crawling.
- Link Analysis.

(4) TEACHING and LEARNING METHODS - EVALUATION

DELIVERY <i>Face-to-face, Distance learning, etc.</i>	Hybrid: Face to face and synchronous distance learning
USE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY <i>Use of ICT in teaching, laboratory education, communication with students</i>	<p>Use of ICT in Teaching During the educational process, various machine learning and programming tools are used, along with the material available at the e-learning platform. The hybrid teaching method involves synchronous learning with the support of the videoconferencing tool Zoom. Students are taught a variety of tools related to the course content and material.</p> <p>Use of ICT in Communication with students</p> <ul style="list-style-type: none"> ● The course material (slides, scientific articles, exercises, etc.) is posted on the course page at the e-learn platform (Moodle).

	<ul style="list-style-type: none"> • Use of Moodle Forums announcements. • Live video meetings via Zoom/Teams. • Contact via email. 																								
<p align="center">TEACHING METHODS</p> <p><i>The manner and methods of teaching are described in detail.</i></p> <p><i>Lectures, recitation, 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.</i></p> <p><i>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</i></p>	<table border="1"> <thead> <tr> <th align="center"><i>Activity</i></th> <th align="center"><i>Semester workload</i></th> </tr> </thead> <tbody> <tr> <td>Lectures</td> <td align="center">30 hrs.</td> </tr> <tr> <td>Recitation</td> <td align="center">9 hrs.</td> </tr> <tr> <td>Exams</td> <td align="center">4 hrs.</td> </tr> <tr> <td>Non-Directed Study</td> <td align="center">107 hrs.</td> </tr> <tr> <td></td> <td></td> </tr> <tr> <td></td> <td></td> </tr> <tr> <td>Course total</td> <td align="center">150 hrs.</td> </tr> </tbody> </table>	<i>Activity</i>	<i>Semester workload</i>	Lectures	30 hrs.	Recitation	9 hrs.	Exams	4 hrs.	Non-Directed Study	107 hrs.					Course total	150 hrs.								
	<i>Activity</i>	<i>Semester workload</i>																							
	Lectures	30 hrs.																							
	Recitation	9 hrs.																							
	Exams	4 hrs.																							
	Non-Directed Study	107 hrs.																							
Course total	150 hrs.																								
<p align="center">COURSE MATERIAL ARRANGEMENT</p>	<p align="center">Theory/Recitation</p> <table border="1"> <tbody> <tr> <td>Overview of Data Structures and Algorithms</td> <td align="center">3 hrs.</td> </tr> <tr> <td>Boolean Retrieval</td> <td align="center">3 hrs.</td> </tr> <tr> <td>The Term Vocabulary and Postings Lists</td> <td align="center">3 hrs.</td> </tr> <tr> <td>Dictionaries and Tolerant Retrieval</td> <td align="center">3 hrs.</td> </tr> <tr> <td>Index Construction</td> <td align="center">3 hrs.</td> </tr> <tr> <td>Index Compression</td> <td align="center">3 hrs.</td> </tr> <tr> <td>Scoring, Term Weighting, and The Vector Space Model</td> <td align="center">3 hrs.</td> </tr> <tr> <td>Computing Scores in a Complete Search System</td> <td align="center">3 hrs.</td> </tr> <tr> <td>Evaluation in Information Retrieval</td> <td align="center">6 hrs.</td> </tr> <tr> <td>Web Search Basics</td> <td align="center">3 hrs.</td> </tr> <tr> <td>Web Crawling and Indexes</td> <td align="center">3 hrs.</td> </tr> <tr> <td>Link Analysis</td> <td align="center">3 hrs.</td> </tr> </tbody> </table>	Overview of Data Structures and Algorithms	3 hrs.	Boolean Retrieval	3 hrs.	The Term Vocabulary and Postings Lists	3 hrs.	Dictionaries and Tolerant Retrieval	3 hrs.	Index Construction	3 hrs.	Index Compression	3 hrs.	Scoring, Term Weighting, and The Vector Space Model	3 hrs.	Computing Scores in a Complete Search System	3 hrs.	Evaluation in Information Retrieval	6 hrs.	Web Search Basics	3 hrs.	Web Crawling and Indexes	3 hrs.	Link Analysis	3 hrs.
	Overview of Data Structures and Algorithms	3 hrs.																							
	Boolean Retrieval	3 hrs.																							
	The Term Vocabulary and Postings Lists	3 hrs.																							
	Dictionaries and Tolerant Retrieval	3 hrs.																							
	Index Construction	3 hrs.																							
	Index Compression	3 hrs.																							
	Scoring, Term Weighting, and The Vector Space Model	3 hrs.																							
	Computing Scores in a Complete Search System	3 hrs.																							
	Evaluation in Information Retrieval	6 hrs.																							
	Web Search Basics	3 hrs.																							
	Web Crawling and Indexes	3 hrs.																							
	Link Analysis	3 hrs.																							
<p align="center">STUDENT PERFORMANCE EVALUATION</p> <p><i>Description of the evaluation procedure</i></p> <p><i>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</i></p> <p><i>Specifically-defined evaluation criteria are given, and if and where they are accessible to students</i></p>	<p>Language of Evaluation: English</p> <p>Evaluation Procedure:</p> <ul style="list-style-type: none"> • Written Midterm (45%). Methods of evaluation: <ul style="list-style-type: none"> ○ Open-ended questions ○ Problem solving • Written Exams (55%). Methods of evaluation: <ul style="list-style-type: none"> ○ Open-ended questions ○ Problem solving <p>The evaluation procedure is announced to the students during the first lecture and is also accessible at the e-learn platform throughout the entire semester.</p>																								
	<p align="center">STUDENT OBLIGATIONS</p> <p><i>Compulsory attendance of lectures, labs, recitations, compulsory participation in midterms, exams, compulsory delivery of homework, projects, etc.</i></p>	<ul style="list-style-type: none"> • Compulsory attendance of lectures • Compulsory attendance of recitation • Compulsory participation in the midterm • Compulsory participation in the exams 																							

(5) ATTACHED BIBLIOGRAPHY

- Suggested Textbooks

1. Manning, Ch. D., Raghavan, P. Schütze, I., Introduction to Information Retrieval, Cambridge University Press, 2008.
2. Baeza-Yates, R. and Ribeiro-Neto, B., Modern Information Retrieval: The Concepts and Technology Behind Search, 2nd Ed., Addison Wesley, 2011.

- Additional Bibliography:

1. Büttcher, S., Clarke, C., Cormack, G., Information Retrieval: Implementing and Evaluating Search Engines, The MIT Press, 2010.
2. Croft, B., Metzler, D., Strohman, T., Search Engines: Information Retrieval in Practice, Addison Wesley, 2010.
3. Witten, I., Moffat, A., Bell, T., Managing Gigabytes: Compressing and Indexing Documents and Images, 2nd Ed., Morgan Kaufmann, 1999.
4. [6] Langville, A.N., and Meyer, C.D., PageRank and Beyond: The Science of Search Engine Rankings, Princeton University Press, 2006.