COURSE OUTLINE

(1) GENERAL

SCHOOL	Science and Technology		
ACADEMIC UNIT	Science and Technology		
PROGRAMME OF STUDIES	MSc in Data Science, MSc in Information and Communication Technology (ICT) Systems, MSc in Mobile and Web Computing: Internet of Things Applications		
LEVEL OF STUDIES	Postgraduate		
COURSE CODE	DSC07, IC11, ME02	SEMESTER	2
COURSE TITLE	Big Data and Cloud Computing		
COURSE TYPE *Elective, compulsory*	Elective, Compulsory, Elective		
INSTRUCTOR(S)	Theory: Prof. Panayiotis Bozanis Lab: Dr. Leonidas Akritidis, Prof. Panayiotis Bozanis		

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*	WEEKLY TEACHING HOURS	CREDITS
	3,4	6
Add rows if necessary. The organisation of teaching and the teaching methods used are described in detail at (d).		

TEACHING ACTIVITIES BREAKDOWN	WEEKLY HOURS
Theory	2,3
Recitation	0
Lab	1,1

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

COURSE TYPE *general background, special background, specialised general knowledge, skills development*	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

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*

On completing the course, the student will be able to:

- Develop the knowledge, understanding and skills to work with Big Data.
- Acquire the necessary algorithmic background to deal with Big Data.
- Apply appropriate analytic techniques and tools to analyzing Big Data.
- Understand Cloud Computing Concepts and Mechanisms.
- Know the concepts, principles, techniques, and methodologies they need to manage cloud services and resources.

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?

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

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

(3) SYLLABUS

The course overviews fundamental principles and concepts of Big Data and Cloud Computing. The students will familiarize themselves with big data and cloud technologies, learn about big data algorithms and cloud computing services.. The topics covered include:

- Big Data Storage and Processing Concepts.
- Hadoop, HDFS, Yarn.
- MapReduce Algorithms.
- Spark.
- NoSQL Databases.
- Cloud Computing Model and Services, Virtualization, Scaling, Capacity and Load Balancing.
- AWS EC2, S3, EMR.

(4) TEACHING and LEARNING METHODS - EVALUATION

DELIVERY *Face-to-face, Distance learning, etc.*	Hybrid: Face to face and synchronous distance learning
USE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY *Use of ICT in teaching, laboratory education, communication with students*	**Use of ICT in Teaching** During the educational process, various big data management and processing platforms 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/platforms related to the course content and material. **Use of ICT in Communication with students**

- The course material (slides, scientific articles, exercises, etc.) is posted on the course page at the e-learn platform (Moodle).
- Use of Moodle Forums announcements.
- Live video meetings via Zoom/Teams.
- Contact via email.

TEACHING METHODS

The manner and methods of teaching are described in detail.
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.

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

Activity	*Semester workload*
Lectures	30 hrs.
Lab	14 hrs.
Homework	12 hrs.
Exams	2 hrs.
Non-Directed Study	92 hrs.
Course total	***150 hrs.***

COURSE MATERIAL ARRANGEMENT

Theory/Recitation

Introduction to Big Data	4 hrs.
Big Data Storage Concepts	2 hrs.
Big Data Processing Concepts: Hadoop, HDFS, MapReduce	4 hrs.
MapReduce Algorithms: Basic Design Techniques	4 hrs.
MapReduce Algorithms: Relational Data, Graph Algorithms	7 hrs.
Big Data Processing Concepts: Spark	4 hrs.
Big Data Storage Technologies - NoSQL Databases	2 hrs.
Cloud Computing: Introduction, Model and Services, Virtualization, Pooling, Capacity Planning, Load Balancing	3 hrs.

Lab

Ubuntu, Hadoop, HDFS	3 hrs.
Yarn, MapReduce	3 hrs.
MapReduce examples	2 hrs.
Spark Examples	2 hrs.
MongoDB, Neo4j	2 hrs.
AWS EC2, S3, EMR	2 hrs.

STUDENT PERFORMANCE EVALUATION

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

Language of Evaluation: English

Evaluation Procedure:

- Written Exams (80%). Methods of evaluation:
  - Open-ended questions.
  - Problem solving.
  - Multiple choice questions (on lab material).
- Homework (20%):
  - Apply Map Reduce and Spark framework to solve Big Data problems.
  - The homework should be completed individually.
  - The students should achieve a passing grade to participate in the written exams.

	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.
STUDENT OBLIGATIONS *Compulsory attendance of lectures, labs, recitations, compulsory participation in midterms, exams, compulsory delivery of homework, projects, etc.*	● Compulsory attendance of lectures ● Compulsory attendance of labs ● Compulsory participation in the exams ● Compulsory delivery of homework

(5) ATTACHED BIBLIOGRAPHY

- Suggested Textbooks

1. Lin, J., Dyer, Ch., Data-Intensive Text Processing with MapReduce, Morgan & Claypool Publishers, 2010.
2. Erl, Th., Khattak, W., Buhler, P., Big Data Fundamentals: Concepts, Drivers & Techniques., Prentice Hall, 2016.
3. Bhowmik, S., Cloud Computing, Cambridge University Press, 2017.
4. Weise,L. Advanced Data Management – For SQL, NoSQL, Cloud and Distributed Databases, De Gruyter Oldenbourg, 2015.
5. White, T. Hadoop: The Definitive Guide, 4th Edition, O'Reilly, 2015.
6. Chambers, B., Zaharia, M., Spark: The Definitive Guide: Big Data Processing Made Simple, O'Reilley, 2018.

- Additional Bibliography:

1. Damji, J.S., Wenig, B., Das, T., Lee, D., Learning Spark: Lightning-Fast Data Analytics, 2nd Edition, O'Reilly, 2020.