

SYLLABUS

1. Data about the program of study

1.1 Institution	The Technical University of Cluj-Napoca
1.2 Faculty	Faculty of Automation and Computer Science
1.3 Department	Computer Science
1.4 Field of study	Computer Science and Information Technology
1.5 Cycle of study	Master
1.6 Program of study/Qualification	Data Science / Master
1.7 Form of education	Full time
1.8 Subject code	1

2. Data about the subject

2.1 Subject name	Introduction to Big Data				
2.2 Course responsible/lecturer	Conf.Dr.Ing. Opreșă Ciprian				
2.3 Teachers in charge of seminars/ laboratory/ project	Șl.Dr.Ing. Joldoș Marius				
2.4 Year of study	1	2.5 Semester	1	2.6 Type of assessment (E - exam, C - colloquium, V - verification)	E
2.7 Subject category	Formative category: DA – advanced, DS – speciality, DC – complementary				DS
	Optionality: DI – imposed, DO – optional (alternative), DF – optional (free choice)				DI

3. Estimated total time

3.1 Number of hours per week	3	of which:	Course	2	Seminars	1	Laboratory	0	Project	0
3.2 Number of hours per semester	52	of which:	Course	28	Seminars	14	Laboratory	0	Project	0
3.3 Individual study:										
(a) Manual, lecture material and notes, bibliography										15
(b) Supplementary study in the library, online and in the field										15
(c) Preparation for seminars/laboratory works, homework, reports, portfolios, essays										15
(d) Tutoring										10
(e) Exams and tests										3
(f) Other activities:										
3.4 Total hours of individual study (suma (3.3(a)...3.3(f)))										58
3.5 Total hours per semester (3.2+3.4)										100
3.6 Number of credit points										4

4. Pre-requisites (where appropriate)

4.1 Curriculum	Artificial Intelligence - bachelor, Distributed systems - bachelor
4.2 Competence	Operating with fundamental computer science concepts

5. Requirements (where appropriate)

5.1. For the course	Blackboard, Projector, PC MS Teams Platform
5.2. For the applications	PC, Specific Software

6. Specific competence

6.1 Professional competences	<p>1 Working with advanced mathematical methods and models, engineering and computing specific techniques and technologies for big data.^[1]</p> <p>2 Development of advanced techniques, methods and methodologies in the domains of software design, programming systems and environments and their applications.^[1]</p> <p>3 Innovative design of artificial intelligence and computer vision systems and related software and hardware using the specific tools.^[1]</p>
------------------------------	---

	<p>4 Contextual integration and exploitation of dedicated information systems.</p> <p>5 Creative pooling of multidisciplinary knowledge in the field of computers and information technology for research, design, optimization, implementation and testing of theories, algorithms and original methods specific to artificial intelligence and computer vision systems.</p> <p>6 Extract information from big data by considering their multivariate nature.</p> <p>7 Design of architectures and solutions for working with big data</p>
6.2 Cross competences	<p>1 Proof of knowledge for the economic, ethical, legal and social context associated with the profession, for correct task identification, schedule of activities, responsible decisions, with the final goal the design, preparation and presentation of a scientific paper.</p> <p>2 Clear and concise description of professional activity flows, tasks and outcomes obtained by assuming the role of leader / project manager or as a member of a research team, as result of personal skills of domain specific information synthesis, global vision, communication skills with collaborators, ability of task stages identification.</p> <p>3 Exercising the skill of continuous self-education and demonstrating critical, innovative and research abilities.</p>

7. Discipline objective (as results from the *key competences gained*)

7.1 General objective	Understanding big data concepts and methods for handling complexities; Solutions identification and design based on a given context; The development of competences and skills for the development of systems for big data
7.2 Specific objectives	Knowledge extraction from big data; Network systems; Ability to design and implement pipelines for big data processing

8. Contents

8.1 Lectures	Hours	Teaching methods	Notes
What is Big Data? Overview and introduction	1	Lectures using blackboard and projector; involving students in debate	
Characteristics and challenges Big Data	1		
Big Data Distributions in Industry. End-to-End Big Data Life cycle overview	2		
Data extraction and transformation. Big data before Apache Spark	2		
Data extraction and transformation. Big data after Apache Spark	2		
Technology stacks. Preprocessing data	2		
Data search and analytics	2		
Technology stacks. Data search and analytics	2		
Data storage	4		
Noise in big data	2		
Extracting knowledge from big data	2		
Social networks and Graph neural networks	2		
Mining data streams	2		
Application: recommender systems	2		
Bibliography			
<ul style="list-style-type: none"> • Thomas Erl, with Wajid Khattak, Paul Buhler Big Data Fundamentals: Concepts, Drivers & Techniques. Prentice Hall. 2016. 978-0134291079 • Ghislain Fourny. The Big Data Textbook - teaching large-scale databases in universities.(2023) https://www.researchgate.net/publication/361334530_The_Big_Data_Textbook_-_teaching_large-scale_databases_in_universities • Jure Leskovec, Anand Rajaraman, Jeff Ullman. Mining of Massive Datasets. http://www.mmids.org/ • Staford - Fundamentals of Big Data https://uit.stanford.edu/service/techtraining/class/fundamentals-big-data-live-online • Stanford - Big Data: Data Engineering at Scale https://uit.stanford.edu/service/techtraining/class/big-data-data-engineering-scale 			

8.2 Applications – Seminars/Laboratory/Project	Hours	Teaching methods	Notes
Tools for processing big data	2	Oral presentation using slides, discussions (Q&A). Using multimedia tools, interactive teaching tools. Using specific software for Big Data	
Hadoop Ecosystem vs No Sql	2		
Processing of big data: MapReduce and Hive for Hadoop	2		
Apache Spark	2		
Data search and analytics: Filtering data, Splunk , Elastic/ELK, Lucene and Apache Solar	4		
Application of knowledge extraction	2		
Bibliography			
<ul style="list-style-type: none"> • Vignesh Prajapati Big Data Analytics with R and Hadoop. Packt Publishing. 2013. 178216328X, 9781782163282 • Peter Ghavami Big Data Analytics Methods. De Gruyter. 2020. 978-1-5474-1795-7 			

*Se vor preciza, după caz: tematica seminariilor, lucrările de laborator, tematica și etapele proiectului.

9. Bridging course contents with the expectations of the representatives of the community, professional associations and employers in the field

The class content was aligned with other similar classes from renowned universities

10. Evaluation

Activity type	Assessment criteria	Assessment methods	Weight in the final grade
Course	Exam	Written exam	70%
Seminar	Exercices. Presentation	Oral examination	30%
Laboratory	-	-	-
Project	-	-	-
Minimum standard of performance: Final grade > 5			

Date of filling in:	Titulari	Titlu Prenume NUME	Semnătura
	Curs	Conf.Dr.Ing. Opreșa Ciprian	
	Aplicații	Șl.Dr.Ing. Joldoș Marius	

Date of approval in the department 20.02.2024	Head of department Prof.dr.ing. Rodica Potolea
Date of approval in the Faculty Council 22.02.2024	Dean Prof.dr.ing. Mihaela Dinsoreanu