SYLLABUS

1. Information regarding the programme

1.1 Higher education institution	**Babeş Bolyai University**
1.2 Faculty	**Faculty of Mathematics and Computer Science**
1.3 Department	**Department of Computer Science**
1.4 Field of study	**Computer Science**
1.5 Study cycle	**Master**
1.6 Study programme / Qualification	**Data Science for Industry and Society**

2. Information regarding the discipline

2.1 Name of the discipline (en) (ro)	**Data Collection an Modeling**					
2.2 Course coordinator	**Lect. Dr. Ing. Grebla Horea Adrian**					
2.3 Seminar coordinator	**Lect. Dr. Ing. Grebla Horea Adrian**					
2.4. Year of study **1**	2.5 Semester	**1**	2.6. Type of evaluation	**E**	2.7 Type of discipline	**Compulsory**
2.8 Code of the discipline	**MME8182**					

3. Total estimated time (hours/semester of didactic activities)

3.1 Hours per week	4	Of which: 3.2 course	2	3.3 seminar/laboratory	1 lab + 1 project
3.4 Total hours in the curriculum	56	Of which: 3.5 course	28	3.6 seminar/laboratory	28
Time allotment:					hours
Learning using manual, course support, bibliography, course notes					36
Additional documentation (in libraries, on electronic platforms, field documentation)					28
Preparation for seminars/labs, homework, papers, portfolios and essays					36
Tutorship					9
Evaluations					10
Other activities:					-

3.7 Total individual study hours	119
3.8 Total hours per semester	175
3.9 Number of ECTS credits	7

4. Prerequisites (if necessary)

4.1. curriculum	•
4.2. competencies	• Average computer usage skills

5. Conditions (if necessary)

5.1. for the course	• Projector
5.2. for the seminar /lab activities	• Computers, Python programming language environment, Excel or Libre Calc

6. Specific competencies acquired

Professional competencies	C3.1 Descrierea de concepte, teorii si modele folosite in domeniul de aplicare C3.3 Utilizarea modelelor si instrumentelor informatice si matematice pentru rezolvarea problemelor specifice domeniului de aplicare C4.5 Încorporarea de modele formale în aplicaţii specifice din diverse domenii
Transversal competencies	CT1. Application of efficient work rules and responsible attitudes towards the scientific domain, for the creative exploitation of one's own potential according to the principles and rules of professional ethics CT2. Efficient conduct of activities organized in an interdisciplinary group and development of empathic capacity of interpersonal communication, networking and collaboration with diverse groups CT3. Use of efficient methods and techniques for learning, information, research and development of abilities for knowledge exploitation, for adapting to the needs of a dynamic society and for communication in a widely used foreign language.

7. Objectives of the discipline (outcome of the acquired competencies)

7.1 General objective of the discipline	• Learn the modes of collecting local and network data using various tools and programming languages and use the basic data structures to store, process and present data
7.2 Specific objective of the discipline	• Study data collection techniques both from local and network (even web) storage and learn the concepts and methods used for data collection. • Study possible simple data models for storing and processing data.

8. Content

8.1 Course	Teaching methods	Remarks

1. Introduction to data (modeling and collection)	• Interactive exposure • Presentation • Explanation • Practical examples • Case-study discussions	
2. File types and the way we can store data in them; data sharing using files		

3. Data Collection
    • Local Data Acquisition
    • Network Data Retrieval		

4. Data Transport/Interchange formats: XML, JSON		
5. Data cleaning		
6. Web crawling		
7. APIs: REST vs SOAP		
8. SQL Basics: modelling		
9. SQL Basics: queries		
10. Simple Data types: list, dictionary, data frames; Python implementation		
11-14. Applications of presented topics using Python		

Bibliography

1. Graeme Simsion Graham Witt, Data Modeling Essentials 3rd Edition, https://www.elsevier.com/books/data-modeling-essentials/simsion/978-0-12-644551-0
2. Steve Hoberman, Data Modeling Made Simple, https://technicspub.com/data-modeling-made-simple/
3. Andy Oppel, Data Modeling, A beginner's guide , https://www.oreilly.com/library/view/data-modeling-a/9780071623988/
4. Marcelo Arenas, Pablo Barceló, Filip Murlak, Foundations of Data Exchange, https://www.cambridge.org/core/books/foundations-of-data-exchange/FABEE52F4A503AA9DEFC54A9382AF54C

8.2 Seminar / laboratory	Teaching methods	Remarks
1. Setup working environment and libraries used (PostgreSQL, Python beautifulsoup4,)	• Interactive exposure • Explanation • Conversation • Didactical demonstration	
2. Working with various file types: csv, XML, JSON, BAM, FASTA		
3. Data storage and retrieval using SQL		
4. Web data crawling		
5. Preparation of individual projects: collect data, clean data, store data.		
6. Project progress and feedback		
7. Project demonstration and presentations		

Bibliography

1. Wes McKinney, Python for Data Analysis, https://wesmckinney.com/pages/book.html
2. SOAP UI, https://www.soapui.org/learn/api/soap-vs-rest-api/
3. Ryan Mitchell, Web Scraping with Python, 2nd edition, https://www.oreilly.com/library/view/web-scraping-with/9781491985564/

9. Corroborating the content of the discipline with the expectations of the epistemic community, professional associations and representative employers within the field of the program

- The course exists in the curriculum of many universities in the world.
- The topics of this course are considered to be a major trend in industry and research.

10. Evaluation

Type of activity	10.1 Evaluation criteria	10.2 Evaluation methods	10.3 Share in the grade (%)
10.4 Course	Demontrate basic knowlwdge of concepts for data collection and modeling	Written exam	50%
10.5 Seminar/lab activities	Apply various techniques for data collection and modeling for a real life problem	Project implementation and presentation	50%
10.6 Minimum performance standards			
Each student should obtain minimum 5 for the written exam and for the final grade.			

Date	Signature of course coordinator	Signature of seminar coordinator
10.05.2022	Lect. dr. ing. Grebla Horea Adrian	Lect. dr. ing. Grebla Horea Adrian

Date of approval   Signature of the head of department

....   Prof. dr. Anca Andreica