COURSE DESCRIPTION

General information		
Lead instructor	Assistant Professor Panče Panov Assistant Professor Andro Milanović	
Course name	**CDS-17: Semantic modeling and data management**	
Study programme	**Computer and Data Science, third cycle Doctoral Study Programme**	
Course status	Optional	
Year	First or Second	
Number of credits and mode of delivery	ECTS student workload coefficient	10
	Number of hours (L+P+S)	30/-/270

Course description

1.1. Course goals

The module contributes to the following general and subject-specific competences

General competences:

- Ability to identify a given research problem, analyse it and offer possible solutions.
- Ability to create new knowledge, which represents a contribution to science.
- Familiarity with the notion of quality and strive for professional quality through autonomy, (self-) criticism, (self-) reflection and (self-) evaluation.

Subject-specific competences:

- Ability to use semantic modeling for representing domain knowledge;
- Ability to model domain knowledge with ontologies;
- Ability to create and use appropriate metadata, that is based on standards;
- Ability to create FAIR resources;
- Ability to plan the data management process
- Ability to manage data both in research and industry

1.2. Course enrolment requirements

There is none.

1.3. Intended course learning outcomes

Knowledge and understanding:

The students:

- Can use different structure for knowledge organization and modeling, use ontologies for semantic modeling by using well established modeling principles and best practices, and use standard schemes for semantic data modelling;
- Can describe types of metadata, recognise metadata formats, identify metadata standards and use metadata standards to describe resources;
- Can paraphrase the FAIR Principles, explain why the FAIR principles were developed, plan for FAIR research outputs and apply the principles to their own work;
- Can describe what a data management plan is, explain why data management planning is a step towards FAIR and sketch a DMP for their own research project;
- Can understand the research data management and industrial data management proces and main use cases, understand the basic principles of Open Research and Open Data, and understand the key roles in data management

1.4. Course content

- Semantic modelling
  - Knowledge organization structures
  - Ontologies as representational artefacts
  - Principles of semantic modelling with ontologies
  - Examples of best practices in semantic modelling
  - Data standardization and examples of schemas for data modelling
- Metadata
  - Types of metadata
  - Metadata formats
  - Metadata standards and their use for describing data resources
- FAIR principles for data management
  - Data life cycle
  - Review of the FAIR data principles
  - Finding and reusing data, data access
  - Data interoperability
  - Data repositories
  - Dealing with confidential, personal, sensitive and private data and ethical aspects
- Data management plan
  - Structure of the data management plan
  - Use of FAIR principles
  - Examples of best practices and tools
- Research data management and data management in industry
  - The process of research data management
  - Open data and open science
  - The process of industrial data management
- Major influence factors and organizational roles in research and industry data management

1.5. Modes of delivery (mark the appropriate boxes with an X)		
	☒ *lectures*	☒ *independent work*
	☐ *seminars and workshops*	☐ *multimedia and network*
	☒ *practicals*	☐ *laboratory*
	☐ *remote learning*	☐ *supervision*
	☐ *field work*	☐ *other* _____

1.6. *Student obligations*

1.7. *Monitoring student work (mark the appropriate boxes with an X)*

Class attendance		Participation in class		Seminar paper		Experimental work	
Written exam		Oral exam		Essay		Research	
Project		Continuous assessment of knowledge		Student report		Practical work	
Portfolio		Schoolwork		Homework			

1.8. *Assessment and evaluation of student work during classes and the final exam*

Type (examination, oral, coursework, project):

- Oral exam, 30%
- Project assignment, 70%

1.9. *Required readings and number of copies relative to the number of students currently taking the course*

Title	*Number of copies*	*Number of students*
Alexopoulos, P. (2020). Semantic modeling for data – Avoiding pitfalls and breaking dilemmas. O'Reilly Media.		
Eryurek, E., Gilad, U., Lakshmanan, V., Kibunguchy-Grant, A., & Ashdown, J. (2021). Data governance: The definitive guide. O'Reilly Media, Inc.		
Strengholt, P. (2023). Data management at scale (2nd ed.). O'Reilly Media, Inc.		
Arp, R., Smith, B., & Spear, A. D. (2015). Building ontologies with basic formal ontology. MIT Press.		
Hedden, H. (2016). The accidental taxonomist (2nd ed.). Information Today, Inc.		
Zozus, M. (2020). The data book – Collection and management of research data. CRC Press.		
Gartner, R. (2016). Metadata – Shaping knowledge from antiquity to the semantic web. Springer.		
Wilkinson, M., Dumontier, M., Aalbersberg, I., … Mons, B. (2016). The FAIR guiding principles for scientific data management and stewardship. Scientific Data, 3, 160018. https://doi.org/10.1038/sdata.2016.18		
Selected state-of-the-art academic papers		

1.10. *Supplementary readings*

1.11. *Methods of quality monitoring that ensure the acquisition of knowledge, skills and competences.*