

UČNI NAČRT PREDMETA / COURSE SYLLABUS**Predmet:** Semantično modeliranje in upravljanje podatkov**Course title:** Semantic modeling and data management

Študijski program in stopnja	Študijska smer	Letnik	Semester
Study programme and level	Study field	Academic year	Semester

Informacijske znanosti, doktorski študijski program tretje stopnje	Računalniške znanosti	Drugi	Tretji ali četrti
Information Sciences, third cycle Doctoral Study Programme	Computer Sciences	Second	Third or fourth

Vrsta predmeta / Course type

Izbirni / Elective

Univerzitetna koda predmeta / University course code:

1-IZ-DR-RZ-IP-SMUP-2024-04-24

Predavanja	Seminar	Vaje	Klinične vaje	Druge oblike študija	Samost. delo	ECTS
Lectures	Seminar	Tutorial	work		Individ. work	
30	-	-	-	-	270	10

Nosilec predmeta / Lecturer: Doc. dr. Panče Panov / Asst. Prof. Panče Panov, PhD**Jeziki / Predavanja / Lectures:** Slovenski / Slovenian, Angleški / English**Languages: Vaje / Tutorial:****Pogoji za vključitev v delo oz. za opravljanje študijskih obveznosti:****Prerequisites:**

Za vključitev v delo ni specifičnih pogojev.

There are no specific requirements for this course.

Vsebina:

1. Semantično modeliranje:
 - a. Strukture za organizacijo znanja;
 - b. Ontologije kot orodje za predstavitev znanja;
 - c. Principi semantčnega modeliranja z ontologijami;
 - d. Primeri dobre prakse pri semantičnem modeliranju;
 - e. Standardizacija podatkov in primeri shem za modeliranje podatkov.
2. Metapodatki:
 - a. Tipi metapodatkov;
 - b. Formati metapodatkov;
 - c. Metapodatkovni standardi in njihova uporaba pri opisovanju podatkovnih virov.
3. Načela FAIR za upravljanje s podatki:
 - a. Življenjski cikel podatkov;
 - b. Pregled načel FAIR;
 - c. Iskanje in ponovna uporaba podatkov, dostop do podatkov;
 - d. Interoperabilnost podatkov;
 - e. Repozitoriji podatkov;
 - f. Ravnanje z zaupnimi, osebnimi, občutljivimi in zasebnimi podatki ter upoštevanje etičnih vidikov.
4. Načrt upravljanja s podatki:
 - a. Struktura načrta upravljanja s podatki;
 - b. Uporaba načel FAIR za pripravo načrta;
 - c. Primeri dobrih praks in orodij.
5. Upravljanje z znanstvenimi in industrijskimi podatki:
 - a. Proces upravljanja z znanstvenimi podatki;
 - b. Odprti podatki in odprta znanost;
 - c. Proces upravljanja s podatki v podjetjih;
 - d. Glavni dejavniki in organizacijske vloge pri upravljanju z

Content (Syllabus outline):

1. Semantic modelling:
 - a. Knowledge organization structures;
 - b. Ontologies as representational artefacts;
 - c. Principles of semantic modelling with ontologies;
 - d. Examples of best practices in semantic modelling;
 - e. Data standardization and examples of schemas for data modelling.
2. Metadata:
 - a. Types of metadata;
 - b. Metadata formats;
 - c. Metadata standards and their use for describing data resources.
3. FAIR principles for data management:
 - a. Data life cycle;
 - b. Review of the FAIR data principles;
 - c. Finding and reusing data, data access;
 - d. Data interoperability;
 - e. Data repositories;
 - f. Dealing with confidential, personal, sensitive and private data and ethical aspects.
4. Data management plan:
 - a. Structure of the data management plan;
 - b. Use of FAIR principles;
 - c. Examples of best practices and tools.
5. Research data management and data management in industry:
 - a. The process of research data management;
 - b. Open data and open science;
 - c. The process of industrial data management;

znanstvenimi in industrijskimi podatki.

d. Major influence factors and organizational roles in research and industry data management.

Temeljni literatura in viri / Readings:

- 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>
-

Cilji in kompetence:

Učna enota prispeva k razvoju naslednjih splošnih in predmetno specifičnih kompetenc:

Splošne kompetence:

- Sposobnost identificiranja danega raziskovalnega problema, njegove analize, ovrednotenja ter oblikovanja možnih rešitev.
- Ustvarjanje novega znanja, ki pomeni relevanten prispevek k razvoju znanosti.
- Prizadevanje za kakovost znanstveno-raziskovalnega dela skozi avtonomnost, (samo)kritičnost, (samo)refleksivnost in (samo)evalviranje.

Predmetno-specifične kompetence:

- Sposobnost uporabe semantičnega modeliranja za predstavitev domenskega znanja;
- Sposobnost modeliranja domenskega znanja z ontologijami;

Objectives and competences:

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

General competences:

- Ability to identify a given research problem, analyse it, evaluate it and formulate possible solutions.
- Ability to create new knowledge, which represents a contribution to science.
- Striving for quality in scientific research through autonomy, (self-)criticism, (self)reflexivity 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;

- Sposobnost ustvarjanja in uporabe metapodatkov, ki temeljijo na standardih;
- Sposobnost ustvarjanja virov, ki temeljijo na FAIR načelih;
- Sposobnost planiranja procesa upravljanja s podatki;
- Sposobnost upravljanja s podatki v raziskovalni dejavnosti in v industriji.

- Ability to manage data both in research and industry.

Predvideni študijski rezultati:

Znanje in razumevanje:

Študenti:

- Znajo izbrati in uporabljati ustrezne strukture za organizacijo znanja in modeliranje ter ontologije za semantično modeliranje z uporabo strokovno priznanih principih in dobrih praks. Prav tako bodo znali kritično oceniti in uporabiti ustrezne standardne sheme za semantično modeliranje podatkov pri svojem raziskovalnem delu.
- Lahko analizirajo in opišejo različne tipe metapodatkov, poznajo različne formate metapodatkov in njihove standarde ter znajo uporabiti te standarde za opisovanje različnih virov.
- Lahko oblikujejo lastne semantične sheme za metapodatke iz svojga raziskovalnega področja.
- Znajo kritično oceniti načela FAIR ter razloge, zakaj so bila razvita ta načela, načrtujejo izsledke raziskav, ki temeljijo na FAIR načelih in uporabijo načela FAIR pri svojem delu.
- Znajo razložiti, kaj je načrt upravljanja s podatki, razložiti, zakaj je načrtovanje en korak k FAIR podatkom.
- Znajo ustvariti načrt upravljanja s podatki za svoj raziskovalni projekt.
- Razumejo procese upravljanja z raziskovalnimi in industrijskimi podatki ter glavne primere uporabe. Prav tako razumejo temeljne principe odprte znanosti in odprtih podatkov ter razumejo glavne vloge pri upravljanju s podatki.

Intended learning outcomes:

Knowledge and understanding:

The students:

- Can choose and use adequate structure for knowledge organization and modeling, use ontologies for semantic modeling by using well established modeling principles and best practices, and critically assess and use standard schemes for semantic data modelling.
- Can analyse and describe types of metadata, recognise metadata formats, identify metadata standards and use metadata standards to describe resources.
- Can design personal semantic metadata schemas for their research area.
- Can critically assess 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.
- Can design 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.

Metode poučevanja in učenja:

- Predavanja z aktivno udeležbo študentov (razlaga, diskusija, vprašanja, primeri, reševanje problemov).
- Individualno projektno delo v obliki priprave raziskovalne projektne naloge.
- Individualne in skupinske konzultacije (diskusija, dodatna razlaga, obrazložitev konkretnih vprašanj); Individualno delo študentov; samostojni študij znanstvene in strokovne literature ter rezultatov raziskav.

Learning and teaching methods:

- Lectures with active participations by the students (explanation, discussion, questions, cases, problems solving).
- Individual seminar work: preparation of research project paper.
- Individual and group consultations (discussions, supplementary explanations, treatment of specific questions).
- Student individual work (literature overview, state-of-the art research in the field).

Delež (v %) /

Načini ocenjevanja:Weight (in %) **Assessment:**

Način (ustni ali pisni izpit, ustno izpraševanje, naloge, projekt):		Type (examination, oral or written, coursework, project):
<ul style="list-style-type: none"> • Ustni izpit 	30 %	<ul style="list-style-type: none"> • Oral exam
<ul style="list-style-type: none"> • Projektna raziskovalna naloga in njena predstavitev. <p>V projektni nalogi bo študent izvedel analizo metapodatkov ter standardnih shem v svojem raziskovalnem področju. Nato bo s semantičnim modeliranjem opisal metapodatke, ki jih uporablja v svojem raziskovalnem projektu. Na koncu bo ustvaril načrt za ravnanje z raziskovalnimi podatki z uporabo načel FAIR.</p>	70 %	<ul style="list-style-type: none"> • Project research paper and its presentation. <p>In the project assignment, the student will conduct an analysis of metadata and standard schemas in their field of research. Then, they will describe the metadata used in their research project through semantic modeling. Finally, they will create a plan for managing research data using FAIR principles.</p>

Reference nosilca / Lecturer's references:

- Kostovska, A., Vermetten, D., Doerr, C., Džeroski, S., Panov, P., & Eftimov, T. (2023) OPTION: OPTImization algorithm benchmarking ONtology. *IEEE transactions on evolutionary computation* 27(6), 1618–1632. <https://dx.doi.org/10.17605/OSF.IO/TD25Y>
- Kostovska, A., Bogatinovski, J., Džeroski, S., Kocev, D., & Panov, P. (2022). A catalogue with semantic annotations makes multilabel datasets FAIR. *Scientific Reports*, 12(1), 7267. <https://doi.org/10.1038/s41598-022-11316-3>
- Kostovska, A., Džeroski, S., & Panov, P. (2020). Semantic description of data mining datasets: An ontology-based annotation schema. In A. Appice, G. Tsoumakas, Y. Manolopoulos, & S. Matwin (Eds.), *Discovery Science. DS 2020. Lecture Notes in Computer Science*, 12323, 140–155. Springer, Cham. https://doi.org/10.1007/978-3-030-61527-7_10

- Vitali, F., Zinno, P., Schifano, E., Gori, A., Costa, A., De Filippo, C., Koroušić Seljak, B., Panov, P., Devirgiliis, C., & Cavalieri, D. (2022). Semantics of dairy fermented foods: A microbiologist's perspective. *Foods*, 11(13), 1939. <https://doi.org/10.3390/foods11131939>
- Panov, P., Soldatova, L. N., & Džeroski, S. (2016). Generic ontology of datatypes. *Information Sciences*, 329, 900–920. <https://doi.org/10.1016/j.ins.2015.08.006>