## UČNI NAČRT PREDMETA / COURSE SYLLABUS

| | |
|---|---|
| **Predmet:** | Podatkovna skladišča in podatkovna analitika |
| **Course title:** | Data warehouses and data analytics |

| Študijski program in stopnja<br>Study programme and level | Študijska smer<br>Study field | Letnik<br>Academic<br>year | Semester<br>Semester |
|---|---|---|---|
| Podatkovne znanosti, magistrski študijski program druge stopnje | - | Drugi | Tretji |
| The second cycle masters study programme Data Sciences | - | Second | Third |

| | |
|---|---|
| **Vrsta predmeta / Course type** | Obvezni / Obligatory |

| | |
|---|---|
| **Univerzitetna koda predmeta / University course code:** | 2-PZ-MAG-PSPA-2020-06-30 |

| Predavanja<br>Lectures | Seminar<br>Seminar | Vaje<br>Tutorial | Klinične<br>vaje<br>work | Druge<br>oblike<br>študija | Samost.<br>delo<br>Individ.<br>work | ECTS |
|---|---|---|---|---|---|---|
| 30 | - | 30 | - | - | 90 | 5 |

| | |
|---|---|
| **Nosilec predmeta / Lecturer:** | izr. prof. dr. Blaž Rodič |

| | | |
|---|---|---|
| **Jeziki /<br>Languages:** | **Predavanja /<br>Lectures:** | slovenski, angleški / Slovene, English |
| | **Vaje / Tutorial:** | slovenski, angleški / Slovene, English |

| **Pogoji za vključitev v delo oz. za opravljanje študijskih obveznosti:** | **Prerequisites:** |
|---|---|
| Predmet zahteva predhodno pridobljena znanja v okviru predmetov Uvod v podatkovno znanost in Vizualizacija podatkov.<br><br>Študent/študentka mora pred pristopom k izpitu imeti pozitivno ocenjene vaje in seminarsko nalogo. | Course requires knowledge obtained within the courses Introduction to Data Science and Data visualization.<br><br>Positively evaluated exercises and seminar paper are a prerequisite for exam. |

| **Vsebina:** | **Content (Syllabus outline):** |
|---|---|
| • Uvod v predmet.<br>Namen študija predmeta, povezanost predmeta z drugimi predmeti, vsebina študija predmeta, študijska literatura.<br>• Poslovni pomen baz podatkov kot ključnega vira znanja za potrebe poslovanja, modeliranje podatkov na različnih ravneh, lastnosti baz podatkov in podatkovnih skladišč. | • Introduction.<br>Aims of the course, relations with other courses, short overview of topics, recommended literature.<br>• Business importance of databases as key sources of knowledge for business purposes, data modelling at various levels, qualities of databases and data warehouses. |

| | |
|---|---|
| • Pregled metod, tehnik in orodij analize podatkov: enostavne rešitve, poslovna analitika, podatkovno rudarjenje, nove tehnologije.<br>• Arhitektura in načrtovanje podatkovnih skladišč: namen in načrtovanje podatkovnih skladišč, zagotavljanje kvalitete shranjenih podatkov, pristopi k analizi shranjenih podatkov.<br>• Priprava, migracija, in posredovanje podatkov, ETL tehnologija, vizualizacija podatkov, kakovost in oplemenitenje podatkov.<br>• Poslovna analitika: OLAP – sprotno analitično procesiranje, odkrivanje, analiza in definiranje poslovnih problemov, kvalitativno in kvantitativno modeliranje, ovrednotenje in prenos rezultatov v poslovno prakso, pregled tipičnih poslovnih problemov in pristopi reševanja.<br>• Veliki podatki (Big Data), Porazdeljene podatkovne tehnologije, orodje Hadoop | • Methods, techniques and tools of data analysis: simple solutions, business analytics, data mining, new technologies.<br>• Data warehouse architecture and warehouse design: design methodologies, data quality assurance, data analysis.<br>• Data preparation, migration and transfer, ETL technology, data visualization, data quality and enhancement.<br>• Business analytics: OLAP - On line analytical processing, business problem detection, analysis, and definition, descriptive and predictive modelling, modelling result evaluation and business adoption, overview of various industry examples with solution approaches<br>• Big Data, Distributed data technologies, orodje Hadoop |

**Temeljni literatura in viri / Readings:**

- Kimball, R., Ross, M. (2013): The Data Warehouse Toolkit: The Definitive Guide to Dimensional Modeling, 3rd Ed., John Wiley & Sons.
- Kimball, R., Ross, M., Thornthwaite, W., Mundy, J., Becker, B. (2016): The Kimball Group Reader: Relentlessly Practical Tools for Data Warehousing and Business Intelligence Remastered Collection, John Wiley & Sons.
- Ballard, C., Farrell, D.M., Gupta, A., Mazuela, C., Vohnik, S. (2012): Dimensional Modeling: In a Business Intelligence Environment, IBM Redbooks.
- Sherman R. (2014): Business Intelligence Guidebook: From Data Integration to Analytics, Morgan Kaufmann.

Izbrana poglavja iz:
- Jaklič, J. (2002): Upravljanje in uporaba podatkov, Ekonomska fakulteta, Ljubljana
- Rob, P. and Coronel, C. (2002): Database Systems: Design, Implementation, and Management, 5th ed., Course Technology, Cambridge MA.
- Giordano, A.D. (2010): Data Integration Blueprint and Modeling: Techniques for a Scalable and Sustainable Architecture, IBM Press.

| **Cilji in kompetence:** | **Objectives and competences:** |
|---|---|
| *Učna enota prispeva k razvoju naslednjih splošnih in predmetno-specifičnih kompetenc:* | *The instructional unit contributes to the development of the following general and subject-specific competences:* |
| *Splošne kompetence:*<br>• Sposobnost skrbeti za kakovost strokovnega dela skozi | *General competences:*<br>• The ability to manage quality of professional work through autonomy, |

avtonomnost, samoiniciativnost, (samo)kritičnost, (samo)refleksivnost in (samo)evalviranje.

*Predmetno-specifične kompetence:*
- Poznavanje arhitekture podatkovnih skladišč.
- Znanje dokumentiranja in analize uporabniških zahtev za razvoj podatkovnih skladišč.
- Znanje priprave sistemskih in funkcijskih specifikacij podatkovnih skladišč.
- Znanje modeliranja podatkovnih skladišč.
- Poznavanje metod integracije podatkov v podatkovnih skladiščih (filtriranje, čiščenje, poenotenje, meta podatki, ETL).
- Poznavanje uporabniških orodij za podatkovno analitiko in poslovno inteligenco.
- Poznavanje zahtev gradnje distribuiranih baz podatkov.
- Poznavanje in osnovna uporaba Hadoop ogrodja za delo z velikimi podatki (Big Data).

initiative, as well as (self-)criticism, (self-)reflection and (self-)evaluation.

*Subject-specific competences:*
- Familiarity with the architecture of data warehouses.
- Knowledge of documentation and analysis of user requirements for the development of data warehouses.
- Skill of preparation of functional specifications, systems and data warehouses.
- Knowledge of data warehouse modelling.
- Familiarity with the methods of integration of data in data warehouses (filtration, cleaning, standardization, metadata, ETL).
- Familiarity with user tools for data analytics and business intelligence.
- Familiarity with the demands of construction of distributed databases.
- Familiarity with and basic use of Hadoop framework for working with Big Data.

| **Predvideni študijski rezultati:** | **Intended learning outcomes:** |
|---|---|
| Znanje in razumevanje:<br><br>*Študent/študentka:*<br>• pozna in razume pomen podatkov in podatkovnih virov;<br>• je zmožen modelirati podatkovna skladišča na konceptualnem, logičnem in fizičnem nivoju;<br>• pozna in razume pomen modeliranja podatkov za potrebe analiz in odkrivanja zakonitosti v podatkih;<br>• pozna in razume vlogo podatkovnih skladišč kot pomembnega sestavnega dela informacijskih sistemov<br>• zna uporabljati tehnike dimenzijskega podatkovnega modeliranja,<br>• razume pomen kakovosti podatkov za uspešno realizacijo podatkovnih skladišč in pozna metodološki | Knowledge and understanding:<br><br>*The student:*<br>• knows and understands the meaning of data and data sources;<br>• knows and understands data models;<br>• can perform data warehouse modelling on conceptual, logical and physical level;<br>• data modelling for the purposes of analysing and discovering rules in data;<br>• knows and understand the role of data warehouses as an integral part of information systems<br>• is able to use the techniques of dimensional data modelling,<br>• understand the importance of data quality for successful implementation of data warehouses, and is familiar with methodological approaches |

| pristop, ki omogoča dvig kakovosti informacij. | that increase the quality of information. |
|---|---|

**Metode poučevanja in učenja:**

- Predavanja z aktivno udeležbo študentov (razlaga, diskusija, vprašanja, primeri, reševanje problemov).
- Laboratorijske vaje (spoznavanje dela s programskimi orodji).
- Individualne in skupinske konzultacije (diskusija, dodatna razlaga, obravnava specifičnih vprašanj).

**Learning and teaching methods:**

- Lectures with active students' involvement (explanation, discussion, questions, examples, problem solving).
- Laboratory work: (using software for data warehouse design and data analytics).
- Individual and group consultations (discussions, supplementary explanations, treatment of specific questions).

| **Načini ocenjevanja:** | Delež (v %) / Weight (in %) | **Assessment:** |
|---|---|---|
| Način (pisni izpit, ustno izpraševanje, naloge, projekt):<br><br>- pisni izpit<br>- empirična seminarska naloga, poročila laboratorijskih vaj | <br><br><br>50 %<br>50 % | Type (examination, oral, coursework, project):<br><br>- written exam<br>- empirical seminar work, report of laboratory exercises |

**Reference nosilca / Lecturer's references:**

- RODIČ, Blaž. Industry 4.0 and the new simulation modelling paradigm. Organizacija : revija za management, informatiko in kadre, ISSN 1318-5454. [Tiskana izd.], aug. 2017, vol. 50, no. 3, str. 193-207, ilustr., doi: 10.1515/orga-2017-0017
- BRELIH, Marjan, RAJKOVIČ, Uroš, RUŽIČ, Tomaž, RODIČ, Blaž, KOZELJ, Daniel. Modelling decision knowledge for the evaluation of water management investment projects. Central European Journal of Operations Research, ISSN 1435-246X, 2019, vol. , iss. , str. https://link.springer.com/content/pdf/10.1007%2Fs10100-018-0600-5.pdf, doi: 10.1007/s10100-018-0600-5.
- KANDUČ, Tadej, RODIČ, Blaž. Optimisation of machine layout using a force generated graph algorithm and simulated annealing. International journal of simulation modelling, ISSN 1726-4529, 2016, vol. 15, no. 2, str. 275-287.
- RODIČ, Blaž, BAGGIA, Alenka. Dynamic airport ground crew scheduling using a heuristic scheduling algorithm. International journal of applied mathematics and informatics, ISSN 2074-1278, 2013, vol. 7, iss. 4, str. 153-163.
- RODIČ, Blaž. Mobile agents for distributed decision support systems. The International Scientific Journal of Management Information Systems, ISSN 1452-774X, 2011, vol. 6, no. 1, str. 20-27.
- RODIČ, Blaž, KLJAJIĆ, Miroljub. Accessing distributed data sources with mobile agents and XML. V: JAŠKOVÁ, Mária (ur.). ECON '05 : [selected research papers], (Research works proceedings, ISSN 0862-7908, Vol. 12, 2005). Ostrava: Technical University of Ostrava, Faculty of Economics. 2005, str. 280-287.

- RODIČ, Blaž, KLJAJIĆ, Miroljub. Integracija simulacijskih orodij v e-poslovni informacijski sistem. V: GRIČAR, Jože (ur.). Izboljšanje konkurenčnosti regije z e-poslovanjem, (Organizacija, ISSN 1318-5454, Letn. 37, 2004, št. 3). Kranj: Moderna organizacija. 2004, str. 162-167.
- ŠKRABA, Andrej, BAGGIA, Alenka, RODIČ, Blaž. Application of a group decision support system in the reform of study programmes. V: DONDON, Philippe (ur.). Recent advances in education and modern educational technologies, (Educational technologies series, 9). [S. l.: s. n.]. 2013, str. 128-134.
- RODIČ, Blaž. Issues of e-collaboration and knowledge management in media industries. V: LUGMAYR, Artur (ur.), et al. Information systems and management in media and entertainment industries, (International series on computer entertainment and media technology (Online), ISSN 2364-9488). Cham: Springer. cop. 2016.