|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **UČNI NAČRT PREDMETA / COURSE SYLLABUS** | | | | | | | | | | | | | | | | | |
| **Predmet:** | | | Sistemi za obdelavo velikih količin podatkov | | | | | | | | | | | | | | |
| **Course title:** | | | Systems for processing large amounts of data | | | | | | | | | | | | | | |
|  | | | | |  | | | | | | | |  | |  | | |
| **Študijski program in stopnja**  **Study programme and level** | | | | | **Študijska smer**  **Study field** | | | | | | | | **Letnik**  **Academic year** | | **Semester**  **Semester** | | |
| doktorski študijski program tretje stopnje Elektrotehnika | | | | | Ni smeri | | | | | | | | 1 | |  | | |
| 3rd cycle: doctoral study programme Electrical Engineering | | | | |  | | | | | | | | **1** | |  | | |
|  | | | | | | | | | | | | | | | | | |
| **Vrsta predmeta / Course type** | | | | | | | | | | | | Izbirni /elective | | | | | |
|  | | | | | | | | | | | |  | | | | | |
| **Univerzitetna koda predmeta / University course code:** | | | | | | | | | | | | 64872 | | | | | |
|  | | | | | | | | | | | | | | | | | |
| **Predavanja**  **Lectures** | **Seminar**  **Seminar** | | | **Vaje**  **Tutorial** | | | **Klinične vaje**  **work** | | | | **Druge oblike študija** | | | **Samost. delo**  **Individ. work** | |  | **ECTS** |
| **15** | **30** | | | **30** | | | **0** | | | | **0** | | | **50** | |  | **5** |
|  | | | | | | | | | | | | | | | | | |
| **Nosilec predmeta / Lecturer:** | | | | | prof. dr. Andrej Kos | | | | | | | | | | | | |
|  | | | | | | | | | | | | | | | | | |
| **Jeziki /**  **Languages:** | | **Predavanja / Lectures:** | | | | Slovenščina ali angleščina / slovene or english. | | | | | | | | | | | |
| **Vaje / Tutorial:** | | | | Slovenščina ali angleščina / slovene or english. | | | | | | | | | | | |
| **Pogoji za vključitev v delo oz. za opravljanje študijskih obveznosti:** | | | | | | | | |  | **Prerequisits:** | | | | | | | |
| Skladno z vpisnimi pogoji. | | | | | | | | |  | In accordance with the entry requirements. | | | | | | | |
| **Vsebina:** | | | | | | | |  | | **Content (Syllabus outline):** | | | | | | | |
| Zbiranje podatkov: pametni telefoni, senzorji in v internet povezane naprave, splet, čiščenje in priprava podatkov, anonimizacija in deidentifikacija podatkov.  Hramba podatkov: razširljive relacijske podatkovne baze, NoSQL podatkovne baze, razumevanje kompromisa med konsistentnostjo podatkov, zmogljivostjo in razpoložljivostjo.  Obdelava podatkov: dogodkovno naravnana obdelava, paralelizacija obdelave (map-reduce), pridobivanje strukturiranih podatkov iz nestrukturiranih.  Analitika: učinkoviti algoritmi za obdelavo in analizo podatkov, strojno učenje.  Vizualizacija: postopki in izzivi vizualizacije velikih količin podatkov, druge modalnosti predstavitve podatkov (soundifikacija, ipd.)  Aplikacije predstavljenih tehnik: sistemi za ugotavljanje konteksta, pametni sistemi (aplikacije pametnih mest, pametnega prometa, ipd.), medicinske aplikacije, socialna omrežja, finančni sistemi | | | | | | | |  | | Data collection: smart phones, sensors and internet-connected devices, web, cleaning and preparation of data, data anonymization and de-identification.  Data retention; scalable relational databases, NoSQL databases, understanding the compromise between the consistency of data, performance and availability.  Data processing: event-oriented processing, processing parallelization (map-reduce), extraction of structured data from unstructured.  Analyses: efficient algorithms for processing and analysis of data, machine learning  Visualization, procedures and challenges of visualizing large amounts of data, other modalities of presentation of data (soundification, etc.).  Applications of the presented techniques: systems for context detection, smart systems (applications of smart cities, smart transport, etc.), medical applications, social networks, financial systems | | | | | | | |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Temeljni literatura in viri / Readings:** | | | | | |
| 1. European Commission: <http://www.internet-of-things-research.eu/pdf/Converging_Technologies_for_Smart_Environments_and_Integrated_Ecosystems_IERC_Book_Open_Access_2013.pdf> 2. Tom White: Hadoop: The Definitive Guide, 3rd Edition; Storage and Analysis at Internet Scale; O'Reilly Media 3. Jure Leskovec, Anand Rajaraman, Jeﬀrey D. Ullman: Mining of Massive Datasets, http://i.stanford.edu/~ullman/mmds/book.pdf 4. Jimmy Lin, Chris Dyer: Data-Intensive Text Processing with MapReduce, http://lintool.github.io/MapReduceAlgorithms/MapReduce-book-final.pdf 5. Tamara Munzner: Visualization Analysis and Design (2014 Draft) http://www.cs.ubc.ca/~tmm/courses/533/book/vispmp-draft.pdf 6. Scott Murray: Interactive Data Visualization for the Web: An Introduction to Designing with D3, O'Reilly Media | | | | | |
| **Cilji in kompetence:** | |  | | **Objectives and competences:** | |
| Pozna pojem »big data«. Zna ovrednotiti količino podatkov, hitrost dogodkov, njihovo raznolikost, ter ključne izzive, povezane z velikimi količinami podatkov.  Pozna razlike, zna izbrati relacijske ali NoSQL podatkovne baze, in ovrednotiti primernost uporabe.  Pozna prednosti in slabosti map-reduce modela ter ovrednotiti v primerjavi z relacijskimi podatkovnimi bazami.  Na primeru zna uporabiti osnovne analitske in vizualizacijske tehnike za delo z velikimi količinami podatkov. | |  | | Is familiar with the concept of "big data". Able to evaluate the amount of data, the rate of events, their diversity, and the key challenges associated with large amounts of data.  Knows the difference and can choose among relational or NoSQL database, and evaluate the appropriateness of use.  Knows the strengths and weaknesses of map-reduce model and evaluates it in comparison with relational databases.  Can apply basic analytical and visualization techniques for working with large amounts of data in a use-case. | |
| **Predvideni študijski rezultati:** | | |  | **Intended learning outcomes:** | |
| Razumevanje pojma »big data«: količina podatkov, hitrost dogodkov, njihova raznolikost, ter ključnih izzivov povezanih z velikimi količinami podatkov.  Razumevanje relacijskih podatkovnih baz, njihovih zmogljivosti in omejitev.  Razumevanje zmogljivosti, prednosti in slabosti NoSQL podatkovnih baz.  Razumevanje map-reduce modela, njegovih prednosti in slabosti, ter primerjave z relacijskimi podatkovnimi bazami.  Razumevanje osnovnih analitskih in vizualizacijskih tehnik za delo z velikimi količinami podatkov. | | |  | Understanding the concept of "big data": data volume, events and their diversity, and key challenges associated with large amounts of data.  Understanding of relational databases, their capabilities and limitations.  Understanding the capabilities, strengths and weaknesses of NoSQL databases.  Understanding of map-reducer model, its strengths and weaknesses, as well as a comparison with relational databases.  Understanding of basic analytical and visualization techniques for working with large amounts of data. | |
|  | | |  |  | |
| **Metode poučevanja in učenja:** | | |  | **Learning and teaching methods:** | |
| Predavanja ali mentorsko delo  Seminar | | |  | Lectures or mentoring  Seminar | |
| **Načini ocenjevanja:** | Delež (v %) /  Weight (in %) | | | | **Assessment:** |
| Oral exam  Seminar  Vaje | 40%  30%  30% | | | | Oral exam  Seminar  Exercise |
| **Reference nosilca / Lecturer's references:** | | | | | |
| KOS, Andrej, PRISTOV, Damijan, SEDLAR, Urban, STERLE, Janez, VOLK, Mojca, VIDONJA, Tomaž, BAJEC, Marko, BOKAL, Drago, BEŠTER, Janez. Open and scalable IoT platform and its applications for real time access line monitoring and alarm correlation. V: 12th International Conference, NEW2AN 2012, and 5th Conference, ruSMART 2012, St. Petersburg, Russia, August 27-29, 2012. ANDREEV, Sergey (ur.), BALANDIN, Sergey (ur.), KOUCHERYAVY, Yevgeni (ur.). Internet of things, smart spaces, and next generation networking : proceedings, (Lecture notes in computer science, ISSN 0302-9743, 7469). Berlin; Heidelberg: Springer, cop. 2012, str. 27-38, ilustr. [COBISS.SI-ID [9370964](http://cobiss.izum.si/scripts/cobiss?command=DISPLAY&base=COBIB&RID=9370964)]  KOS, Andrej, SEDLAR, Urban, STERLE, Janez, VOLK, Mojca, BEŠTER, Janez, BAJEC, Marko. Network monitoring applications based on IoT system. V: Proceedings of the 2013 18th European Conference on Network and Optical Communications & 2013 8th Conference on Optical Cabling and Infrastructure (NOC-OC&I). Graz: University of Technology, Institute of Microwave and Photonic Engineering, 2013, str. 69-73, ilustr. [COBISS.SI-ID [9963860](http://cobiss.izum.si/scripts/cobiss?command=DISPLAY&base=COBIB&RID=9963860)]  SEDLAR, Urban, VOLK, Mojca, STERLE, Janez, SERNEC, Radovan, KOS, Andrej. Contextualized monitoring and root cause discovery in IPTV systems using data visualization. IEEE network, ISSN 0890-8044, Nov.-Dec. 2012, vol. 26, no. 6, str. 40-46, ilustr. [COBISS.SI-ID [9594452](http://cobiss.izum.si/scripts/cobiss?command=DISPLAY&base=COBIB&RID=9594452)]  PETERNEL, Klemen, POGAČNIK, Matevž, TAVČAR, Rudi, KOS, Andrej. A presence-based context-aware chronic stress recognition system. Sensors, ISSN 1424-8220, Nov. 2012, vol. 12, no. 11, str. 15888-15906, ilustr. http://www.mdpi.com/1424-8220/12/11/15888, doi: 10.3390/s121115888. [COBISS.SI-ID 9534292]  KOS, Andrej, SEDLAR, Urban, PETERNEL, Klemen, VOLK, Mojca, STERLE, Janez, ZEBEC, Luka, VIDONJA, Tomaž, BEŠTER, Janez. Odprta komunikacijska platforma IoT. V: Petindvajseta delavnica o telekomunikacijah, 12. in 13. maj 2011, Brdo pri Kranju. SIMIČ, Nikolaj (ur.), et al. Internet stvari : zbornik referatov, (VITEL, ISSN 1581-6737). Ljubljana: Elektrotehniška zveza Slovenije, cop. 2011, f. 11-15, ilustr. [COBISS.SI-ID[8409684](http://cobiss.izum.si/scripts/cobiss?command=DISPLAY&base=COBIB&RID=8409684)] | | | | | |