

Področje dela Odseka za tehnologije znanja je razvoj naprednih informacijskih tehnologij za zajemanje, shranjevanje in upravljanje znanja, ki so praktično uporabne za razvoj informacijske in na znanju temelječe družbe. Uveljavljena področja našega dela vključujejo inteligentno analizo podatkov, besedil in spleta (strojno učenje, rudarjenje podatkov, odkrivanje zakonitosti v podatkih), jezikovne tehnologije in računalniško jezikoslovje, računalniško kreativnost, podporo pri odločanju ter upravljanje znanja. Razvijamo tudi aplikacije teh tehnologij na področjih znanosti in upravljanja z okoljem, agronomije, medicine, biomedicine in bioinformatike, ekonomije, financ in trženja. Odsek je tudi uveljavljeno središče jezikoslovja in digitalne humanistike.



Vodja:

prof. dr. Nada Lavrač

Leta 2017 smo bili vključeni v štirinajst domačih projektov, dva evropska projekta 7. OP in deset projektov Obzorja 2020, eno COST-akcijo, en INTERREG V-A Italija-Slovenija, en infrastrukturni projekt, en projekt pametne specializacije, dva bilateralna projekta ter pet tržnih projektov. Na odseku se je usposabljal osem mladih raziskovalcev.

Na področju **inteligentne analize in rudarjenja podatkov** smo razvili več novih metod in jih uporabili v raznovrstnih aplikacijah. Na področju analize heterogenih omrežij smo razvili nove hevrstike preiskovanja, ki temeljijo na hevrstikah tekstovnega rudarjenja. Na področju semantičnega rudarjenja podatkov smo v postopku predprocesiranja podatkov z metodami analize omrežij dosegli stokratno pospešitev preiskovanja prostora semantičnih opisov. Na področju učenja poopisov smo razvili metodo za generiranje množic poopisov z večciljnim napovednim razvrščanjem ter nov okvir in metodo za generiranje množic opisnih pravil. Objavili smo metode, ki omogočajo analizo velikih podatkov v okviru spletne platforme ClowdFlows, katere nadaljnji razvoj ter možnost komercializacije platforme financira evropski projekt FET Innovation Launchpad projekt CF-Web. V okviru projekta PD-manager, ki ga koordiniramo na IJS, smo razvili metodo, ki kombinira večciljno učenje in analizo kratkih časovnih vrst za identifikacijo skupin podobnih pacientov s parkinsonovo boleznijo. Začeli smo delo pri novem projektu H2020 SAAM (Supporting Active Ageing through Multimodal coaching), kjer razvijamo sistem virtualnega trenerja za podporo starejšim osebam, ki živijo doma.

Razvili smo nov način samodejnega modeliranja dinamičnih sistemov, ki upošteva tako podatke kot domensko predznanje. Gre za numerično optimizacijo, ki temelji na nadomestnih modelih. Ti omogočajo učinkovitejšo optimizacijo v diferencialnih enačbah, ki modelirajo dani dinamični sistem.

Uspešno smo končali projekt MAESTRA (Learning from Massive, Incompletely Annotated, and Structured Data), ki so mu recenzenti iz EU podelili najvišje ocene. Srž projekta je bila analiza kompleksnih podatkov, kar vključuje tudi polnadzorovano učenje z delno označenimi podatki.

V okviru projekta smo razvili številne nove metode za napovedovanje različnih vrst strukturiranih vrednosti. Metode delujejo tako v paketnem načinu učenja (polnadzorovana klasifikacija, večciljna regresija) kot pri učenju iz podatkovnih tokov (drevesne metode za večciljno regresijo), opcijska drevesa za hierarhično večoznačno klasifikacijo, drevesa za hierarhično večciljno regresijo in ansambli dreves z naključnim izbiranjem oznak za večoznačno klasifikacijo. Razvili smo tudi metode za rangiranje značilk za različne vrste strukturiranih vrednosti (večoznačna regresija in hierarhična večoznačna klasifikacija).

Razvite metode smo uporabili pri analizi slik in modeliranju rasti lesne zaloge v gozdu. Na področju medicine smo metode uporabili za modeliranje časovnih vrst meritev glukoze pri sladkornih bolnikih in analizo vpliva patogenov na mikrobioto v človeškem črevesju. Metode smo uporabili tudi za napovedovanje porabe 33 električnih grelnikov satelita Mars Express Evropske vesoljske agencije.

V okviru projekta HBP (Human Brain Project, FET Flagship) razvijamo nove metode za rudarjenje podatkov in jih uporabljamo za odkrivanje bioloških podpisov nevrodegenerativnih bolezni, kot je alzheimerjeva bolezen. Razviti sta bili novi metodi za rudarjenje večopisnih množic. Obe temeljita na napovednem razvrščanju, uporabili pa smo ju pri iskanju povezav med kliničnimi in biološkimi kazalci alzheimerjevih bolnikov s posebnimi potrebami.

V okviru projekta LANDMARK smo nadaljevali razvoj modelov za napovedovanje funkcij tal. Večina naše raziskovalne dejavnosti je bila osredinjena na predpripravo podatkov, kar je bila zelo zapletena naloga. Na osnovi

Nada Lavrač je imela vabljen predavanje na 16. mednarodni konferenci semantičnega spleta ISWC-2017 na Dunaju z naslovom »Od relacijskega do semantičnega podatkovnega rudarjenja«. Konferenca je osrednji dogodek o semantičnem spletu v svetovnem merilu z več kot 600 udeleženci.

Martin Breskvar, Matej Petkovič in Blaž Škrli so bili del zmagovalne ekipe na hackatlonu HackElect 2017. Z metodami, kot so napovedna klasifikacijska drevesa in globoke nevronske mreže, je edinim uspelo napovedati porabo elektrike na nivoju posamezne gospodinjke naprave za različne časovne skale.

model najprej validirala in nato s kalibracijo priredila za njihovo uporabo v treh prevladujočih pedoklimatskih conah Evrope. Podatke, pridobljene iz različnih nacionalnih podatkovnih baz in pristojnih EU-institucij, smo kritično prevzeli in izvedli njihovo pripravo za uporabo v razvitih odločitvenih modelih. To je zahtevalo veliko usklajevanja med različnimi partnerji, ki so prispevali podatke, pri čemer je bila glavnina dela določitev kritičnih vrednosti za njihovo diskretizacijo.

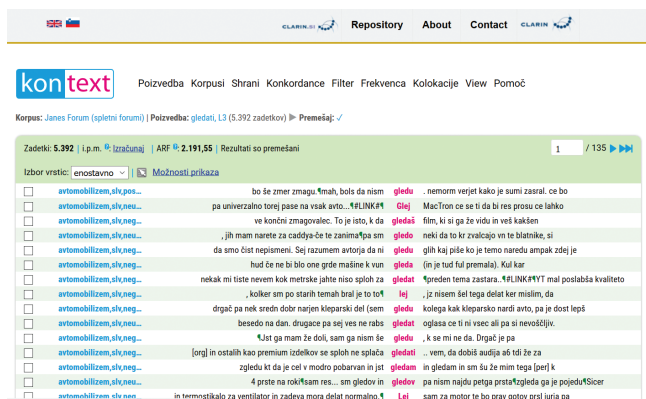
Velik del naše dejavnosti pri izvajanju projekta BioDiv je bila predpriprava podatkov. Dokončana podatkovna baza je sestavljena iz treh velikih skupin atributov, ki opisujejo: i) taksonomske in funkcionalne lastnosti žuželk iz družine Syrphidae, sirfid, ii) lastnosti pokrajine, ki opisujejo strukturne značilnosti polj in naravnih habitatov, ki jih obdajajo, in iii) klimatske razmere v obdobju, ko je potekalo vzorčenje vrst žuželk iz družine Syrphidae. Številčnost najpogostejših vrst Syrphidae in njihovih funkcionalnih skupin smo modelirali z uporabo modelskih in regresijskih dreves.

V letu 2017 smo začeli izvajati dva evropska projekta H2020: TRUE (TRAnsition paths to sUstainable legume based systems in Europe) in TomRes (A novel and integrated approach to increase multiple and combined stress tolerance in plants using tomato as a model). Ker je večina naše aktivnosti načrtovana v letih 2018 in 2019, smo se v letu 2017 osredinili predvsem na seznanjanje z raziskovalnimi vsebinami obeh projektov. Pri tem smo se posvetili intenzivnemu pregledu objav na našem vsebinskem področju raziskav ter sestavili prve osnutke konceptualnih modelov za oceno trajnosti kvalitetnih verig (TRUE) oz. modelov za napovedovanje ekofizioloških odzivov paradiznika na stres, povzročen s pomanjkanjem vode in hranil (TomRes).

Na področju analize besedil, spletnih podatkov in heterogenih informacijskih omrežij smo uspešno končali projekt SIMPOL (Financial Systems Simulation and Policy Modelling). Nadaljevali smo delo pri FET-projektu Obzorja 2020 DOLFINS (Distributed Global Financial Systems for Society). Nadaljevali smo tudi delo pri nacionalnem projektu FORMICA (Vpliv formalnih in neformalnih komunikacij na trg kapitala). Glavni poudarki raziskav na tem področju so v povezavi rudarjenja besedil, analize omrežij in analize sentimenta s ciljem poiskati zanimive lastnosti in povezave v različnih domenah. Glavni viri podatkov, ki jih analiziramo, so družbena omrežja (Twitter in Facebook), spletne novice, letna poročila podjetij in drugi relevantni podatki (npr. glasovanje na referendumu o brexitu, cene delnic, register lobistov itd.).

Razvili smo metodologijo za ustrezno evalvacijo modelov za analizo sentimenta (oz. odnosa) na Twitterju, kjer so podatki časovno urejeni. Klasifikacijski modeli so specifični in odvisni od uporabljenega jezika in tematike. Naš način temelji na obsežni množici ročno označenih tvitov, ki so jih s sentimentom oz. odnosom označili strokovnjaki za izbrano tematiko. Z metodami strojnega učenja smo nato zgradili visokokvalitetne tematsko in

jezikovno odvisne klasifikacijske modele. Pokazali smo, da je navaden način evalvacije modelov s prečno validacijo primeren le, če uporabimo bločno prečno validacijo brez naključnega mešanja učnih primerov. Zgoraj opisano metodologijo smo uporabili na zelo relevantnem evropskem primeru: na referendumu o izstopu Velike Britanije iz EU (brexit). Razvili smo specializiran model, ki klasificira tvite glede na pričakovani odnos avtorja do brexita (za, proti ali neodločen). Spremljali smo mnenje javnosti v VB pred in med referendumom. Prav tako smo identificirali najvplivnejše podpornike in nasprotnike brexita na Twitterju in ugotovili, da so bili podporniki brexita občutno aktivnejši od nasprotnikov. Poleg tega smo natančno analizirali količino tvitov in njihov odnos glede na gibanje 30 podjetij v indeksu Dow Jones Industrial Average (DJIA) v obdobju treh let. Osredinili smo se na četrtletna poročila podjetij in pokazali, da je velika razlika med natančnim časom objave poročil: preden se trgovanje začne ali po koncu dnevnega trgovanja. Analizirali smo razlike glede na količino tvitov, kumulativne nenormalne donose, dejanske donose in nepričakovane donose. Zbrali in analizirali smo tudi letna poročila podjetij v indeksu DJIA ter preučili korelacije med jezikovnimi in finančnimi indikatorji.



The screenshot shows the KonText web application interface. At the top, there is a navigation bar with links for Repository, About, and Contact. Below this, the main header displays 'konText' and 'Poizvedba Korpusi Shrani Konkordance Filter Frekvenca Kolokacije View Pomoč'. The main content area shows search results for the word 'gledati' in the Janes Forum corpus. It includes a search bar with the text 'Zadeki: 5.392 | 1 p.m. | Iščite: ARF 2.191,55 | Rezultati so premešani'. Below the search bar, there is a table of search results with columns for 'Izbor vrstic', 'gledati', and 'gledati'. The table contains several rows of search results, each with a checkbox, the word 'gledati', and a frequency count. The interface is designed to allow users to search for specific words in a corpus and view the results in a structured format.

Slika 1: Konkordance glagola "gledati" v korpusu Janes Forum. Spletni konkordancnik KonTeks raziskovalne infrastrukture CLARIN.SI ponuja več kot 40 korpusov, med njimi več novih korpusov uporabniško generirane slovenščine. Poizvedba vrne tudi nestandardne oblike glagola, na primer "gledo".

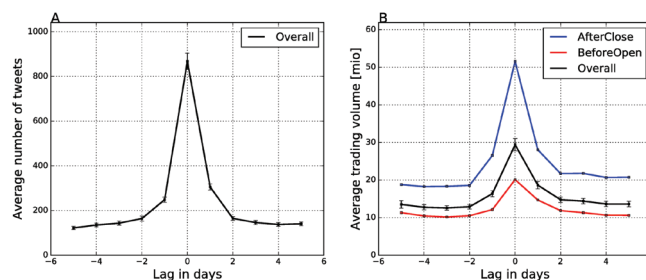
Dolgoročni cilj na področju **podpore pri odločanju** je razvijati metode in tehnike odločitvenega modeliranja, jih podpreti z računalniškimi orodji ter jih povezovati s sistemi za rudarjenje podatkov. V letu 2017 smo objavili metodološki članek, v katerem poleg formalnega opisa naše najbolj uporabljane metode za odločitveno modeliranje DEX predlagamo številne razširitve načina glede uporabe numeričnih odločitvenih atributov, porazdelitev vrednosti in relacijskega vrednotenja alternativ. Objavili smo tudi rezultate uporabe teh metod pri dolgoročni presoji vzdržne proizvodnje električne energije v Sloveniji. Nadaljevali smo delo pri dveh EU-projektih H2020, ki sta namenjena razvoju računalniških platform in sistemov za pomoč pri odločanju pri upravljanju težkih kroničnih bolezni: parkinsonove bolezni (projekt PD_manager) in popuščanja srca (HeartMan). Raziskali smo uporabnost treh metod za aproksimacijo funkcij koristnosti pri metodi DEX, analizirali statistične lastnosti modelov DEX in opravili začetne preizkuse uvajanja povratnih zank (ciklov) v modele DEX.

Na področju **jezikovnih tehnologij in digitalne humanistike** se ukvarjamo z izdelavo jezikovnih virov in metod za jezikoslovno označevanje besedil, pri čemer se osredinjamo predvsem na slovenski jezik. V letu 2017 smo sodelovali pri izdelavi tekstnokratične izdaje Poljanskega rokopisa iz leta 1800, ki obsega 700 strani, pri razvoju univerzalne odvisnostne drevesnice za srbsčino, pri organizaciji evalvacijske kampanje VarDiala (4th Workshop on NLP for Similar Languages, Varieties and Dialects) in pri razširitvi referenčnega korpusa slovenščine Gigafida s spletnimi besedili. Opisali smo postopke in orodja, potrebna za razvoj smernic pri jezikoslovnem označevanju, za gradnjo večjezičnih jezikovnih virov MULTEXT-East in za referenčni oblikoslovni korpus slovenščine Sloleks ter podrobno opisali njegov nastanek.

Končevali smo delo pri nacionalnem raziskovalnem projektu JANES »Viri, orodja in metode za jezikoslovno analizo nestandardne slovenščine«, kjer smo prispevali pregled najboljših praks za projekte, ki se ukvarjajo z računalniško posredovano komunikacijo, analizirali nestandardni jezik v tvitih in razvili način za obliko-skladenjsko označevanje nestandardne slovenščine, hrvaščine in srbsčine. Razvili smo več načinov za napovedovanje spola avtorjev v družbenih medijih: jezikovno neodvisen sistem za napovedovanje na podlagi nebesedilnih informacij, sistem za razpoznavanje spola avtorjev in jezikovne različice iz besedil, s katerim smo dosegli 2. mesto na tekmovanju PAN, ter sistem za napovedovanje spola, ki se osredinja na slovenske tvite. Zgradili smo dva ročno označena učna korpusa nestandardne slovenščine, tj. skladenjsko označeni Janes-Syn ter drugo različico korpusa Janes-Tag, v kateri so k normaliziranim besednim oblikam dodane oznake imenskih entitet, obliko-skladenjske oznake in leme. Izdelali smo tudi učne množice za nestandardno hrvaščino in srbsčino, ki vsebujejo enake ravni označevanja kot Janes-Tag. Sestavili smo tudi ročno označene učne množice za jezikoslovne raziskave nestandardnega jezika, in sicer Janes-Vejica, ki vsebuje oznake (ne)pravilno postavljenih vejic, Janes-Kratko, v katerem so označeni pojavi krajšanja sporočil na Twitterju, in Janes-Preklop, ki označuje primere kodnega preklapljanja. Ob koncu projekta smo dokončali glavni prispevek projekta, tj. različico 1.0 korpusa JANES, ki vsebuje 250 milijonov pojavnic, vsebovanih v 13 milijonih besedil spletnih uporabniških vsebin v slovenskem jeziku, ki so jim dodani obširni metapodatki, obliko-skladenjske oznake in leme. Po celotnem korpusu je mogoče prosto iskati na spletu prek konkordančnikov, ki jih ponuja infrastruktura CLARIN.SI, prek repozitorija CLARIN.SI pa je mogoče naložiti s spleta posamezne podkorpuse: Janes-Tweet, Janes-Blog, Janes-Forum, Janes-News in Janes-Wiki.

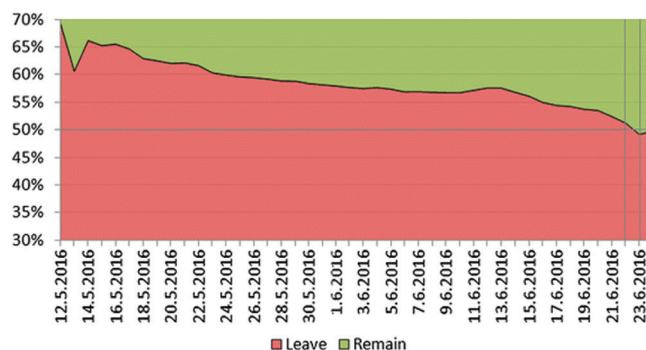
Nadaljevali smo delo pri nacionalnem raziskovalnem projektu KAS »Slovenska znanstvena besedila: viri in opis«, kjer smo analizirali leksiko že zgrajenega korpusa KAS, ki vsebuje skoraj milijardo pojavnic iz znanstvenih besedil, zbranih v digitalnih knjižnicah slovenskih univerz. Začeli smo delo pri našem novem nacionalnem projektu FRENK »Viri, metode in orodja za razumevanje, prepoznavanje in razvrščanje različnih oblik družbeno nesprejemljivega diskurza v informacijski družbi«, kjer smo razvili shemo označevanja družbeno nesprejemljivih oblik diskurza na spletu.

V okviru industrijskega projekta TermIolar smo za prevajalsko podjetje Iolar izdelali sistem za polavtomatsko urejanje terminologije. Preučili smo



Slika 2: Dnevno število tvitov (A) in obseg trgovanja (B) v času okoli izdaje obvestil o zaslučkih (OZ). Skupno povprečno število tvitov na trgovni dan je 200. Obseg trgovanja (B) prikazuje skupno povprečje v vseh OZ (črna črta), povprečno trgovanje za AfterClose (modra črta) in obvestila o predhodnem objavljanju (rdeča črta). Okrog črnih črt je označena velikost ene standardne napake.

Nikola Ljubešić in Yves Scherrer (Univerza v Ženevi) sta zmagala pri reševanju Odprte naloge CLIN2017 o normalizaciji zgodovinskih besedil z orodjem CSMTiser, ki je bilo razvito v okviru nacionalnega temeljnega raziskovalnega projekta JANES in nacionalne raziskovalne infrastrukture CLARIN.SI.



Slika 3: Kumulativno stališče uporabnikov Twitterja v povezavi z brexitom. Rdeča barva predstavlja uporabnike, ki so za brexit, zelena barva pa tiste, ki so proti (neutrálni uporabniki niso vključeni). Uporabniki, ki so bili proti, so se postopoma pridruževali, v nasprotju z uporabniki, ki so bili za - ti so bili večinoma prisotni že dalj časa in so prevladovali v številu tvitov. Šele 2 dni pred referendumom je število uporabnikov, ki so bili za, preseglo število uporabnikov, ki so bili proti.

načine za luščenje dobrih primerov rabe ter urejanje pomnilnikov prevodov z razvrščanjem besedil na podlagi ontologij.

Upravljamo slovensko raziskovalno infrastrukturo CLARIN.SI, ki raziskovalcem s področja humanistike in družbenih ved omogoča preprosto objavo in trajen dostop do digitalnih jezikovnih podatkov. V letu 2017 se je število vnosov v repozitorij CLARIN.SI skoraj podvojilo, konec leta je bilo v njem 79 vnosov. Sodelovali smo pri izdelavi druge različice govornega korpusa Gos VideoLectures pri ročno označenem korpusu ssj500k in pri korpusu razprav v slovenskem parlamentu SlovParl. V sklopu infrastrukture CLARIN.SI smo poskrbeli za delovanje dveh konkordančnikov, KonText and noSketch Engine, in s tem omogočili iskanje po več kot 40 korpusih ter njihovo analizo. Nadaljevali smo tudi integracijo orodij za označevanje slovenščine v platformo za zasnovo in uporabo spletnih delotokov WebLicht. CLARIN.SI je podprl tudi organizacijo poletne šole prevajalskih tehnologij »TransTech17: 3rd Summer School in Translation Technologies« (4.–8. september 2017) in seminarja o uporabi empiričnih podatkov v jezikoslovju »ReLDI Seminar: Empirical data in linguistics: From research design to data analysis« (21.–23. junij 2017).

Sodelovali smo tudi s Slovenskim inštitutom za standardizacijo. Kot slovenski predstavniki v ISO/TC37/SC4 (Terminology and Other Language and Content Resources / Language Resources Management) smo bili odgovorni za ocenjevanje, prevajanje in potrjevanje slovenskih standardov na tem področju. Še naprej delujemo tudi kot tehnični uredniki pri spletni različici Slovenskega biografskega leksikona.

Najpomembnejše objave v preteklem letu

1. Pavle Boškosi, Andrej Debenjak, Biljana Mileva-Boshkoska. Rayleigh copula for describing impedance data - with application to condition monitoring of proton exchange membrane fuel cells. *European journal of operational research*, ISSN 0377-2217. [v tisku] 2017, 21 pp., doi: 10.1016/j.ejor.2017.08.058
2. Jurica Levatić, Michelangelo Ceci, Dragi Koccev, Sašo Džeroski. Self-training for multi-target regression with tree ensembles. *Knowledge-based systems*, ISSN 0950-7051. 123 (2017), 41–60, doi: 10.1016/j.knsys.2017.02.014
3. Manja Klemenčič, Martin Žnidaršič, Anže Vavpetič, Matej Martinc. Erasmus students' involvement in quality enhancement of Erasmus+ mobility through digital ethnography and ErasmusShouts. *Studies in higher education*, ISSN 0307-5079, 42 (2017) 5, 925–932, doi: 10.1080/03075079.2017.1293879
4. Tomaž Erjavec. MULTEXT-East. V: Nancy M. Ide, James Pustejovsky (ur.). *Handbook of linguistic annotation*. Amsterdam: Springer, (2017), 441–462
5. Grčar, Miha, Čerepnalkoski, Darko, Mozetič, Igor, Kralj Novak, Petra. Stance and influence of Twitter users regarding the Brexit referendum. *Computational social networks*, ISSN 2197-4314, 4 (2017), 6-1–6-25, doi: 10.1186/s40649-017-0042-6
6. Peter Gabrovšek, Darko Aleksovski, Igor Mozetič, Miha Grčar. Twitter sentiment around the earnings announcement events. *PloS one*, ISSN 1932-6203, 12 (2017) 2, e0173151-1–e0173151-21, doi: 10.1371/journal.pone.0173151
7. Marko Bohanec, Nejc Trdin, Branko Kontić. A qualitative multi-criteria modelling approach to the assessment of electric energy production technologies in Slovenia. *Central European Journal of Operations Research*, ISSN 1435-246X, 25 (2017) 3, 611–625, doi: 10.1007/s10100-016-0457-4
8. Matej Mihelčič, Marko Bohanec. Approximating incompletely defined utility functions of qualitative multi-criteria modeling method DEX. *Central European Journal of Operations Research*, ISSN 1435-246X, 25 (2017) 3, 627–649, doi: 10.1007/s10100-016-0451-x
9. Janez Kranjc, Roman Orač, Vid Podpečan, Nada Lavrač, Marko Robnik Šikonja. ClowdFlows : online workflows for distributed big data mining. *FGCS*, ISSN 0167-739X. 68 (2017), 38–58, doi: 10.1016/j.future.2016.07.018
10. Matej Mihelčič, Sašo Džeroski, Nada Lavrač, Tomislav Šmuc. A framework for redescription set construction. *Expert systems with applications*, ISSN 0957-4174. 68 (2017), 196–215, doi: 10.1016/j.eswa.2016.10.012
11. Sabina Horvat, Aleksander Mahnič, Martin Breskvar, Sašo Džeroski, Maja Rupnik. Evaluating the effect of *Clostridium difficile* conditioned medium on fecal microbiota community structure. *Scientific reports*, ISSN 2045-2322, 17 (2017), 11, doi: 10.1038/s41598-017-15434-1.
12. Dragan Gamberger, Nada Lavrač, et al. Identification of clusters of rapid and slow decliners among subjects at risk for Alzheimer's disease. *Scientific reports*, ISSN 2045-2322, 7 (2017), 1–12, doi: 10.1038/s41598-017-06624-y

Organizacija konferenc, kongresov in srečanj

1. The European Conference on Machine Learning and Principles and Practice of Knowledge Discovery – ECML PKDD 2017, Skopje, 18.–22. 9. 2017
2. Organizacija sestanka EU projekta MAESTRA, Ljubljana, 3.–7. 4. 2017

3. Organizacija sestanka podprojekta »Medical Informatics Platform (SP8)«, ki se izvaja v sklopu EU-projekta HBP, Ljubljana, 25.–27. 1. 2017
4. "The Tenth International Ljubljana-Zagreb Workshop on Knowledge Technologies", Fiesa, 15.–16. 6. 2017

Nagrade in priznanja

1. Marko Bohanec: Fellow and Distinguished Scholar, International Institute for Applied Knowledge Management, 2017
2. Martin Breskvar, Matej Petkovič in Blaž Škrlj, hackatlon HackElect 2017 (zmagovalna ekipa)
3. Nikola Ljubešić, Yves Scherrer (Univerza v Ženevi) sta zmagala pri reševanju Odprte naloge CLIN2017 o normalizaciji zgodovinskih besedil z orodjem CSMTiser, ki je bilo razvito v okviru nacionalnega temeljnega raziskovalnega projekta JANES in nacionalne raziskovalne infrastrukture CLARIN.SI.
4. Matej, Martinc, Iza Škrjanec, Katja Zupan, Senja Pollak: 2. mesto na tekmovanju PAN 2017 (PAN2017: profiliranje avtorjev - napovedovanje spola in jezikone različice) na CLEF 2017 - Conference and Labs of the Evaluation Forum, Dublin, Irska.
5. Nikola Ljubešić, Yves Scherrer (Univerza v Ženevi) sta zmagala pri reševanju Odprte naloge CLIN2017 o normalizaciji zgodovinskih besedil z orodjem CSMTiser, ki je bilo razvito v okviru nacionalnega temeljnega raziskovalnega projekta JANES in nacionalne raziskovalne infrastrukture CLARIN.SI.

MEDNARODNI PROJEKTI

1. BIODIV - Razumevanje in upravljanje biološke raznovrstnosti agrarnih ekosistemov s pomočjo podatkovnega rudarjenja in odločitvenega modeliranja; Napovedovanje strukturiranih vrednosti z uporabo v trajnostni kmetijski proizvodnji
Arvalis - Institut Du Végétal
prof. dr. Sašo Džeroski
2. 7. OP - SIMPOL; Simulacija finančnih sistemov in modeliranje politik
Evropska komisija
prof. dr. Igor Mozetič
3. 7. OP - MAESTRA; Učenje iz obsežnih, nepopolno označenih in strukturiranih podatkov
Evropska komisija
prof. dr. Sašo Džeroski
4. PARSEME: Razčlenjevanje in večbesedni izrazi. Jezikoslovna natančnost in računalniška učinkovitost pri procesiranju naravnih jezikov
Cost Office
prof. dr. Tomaž Erjavec
5. H2020 - DOLFINS; Distribuirani globalni finančni sistemi za družbeno dobro
Evropska komisija
prof. dr. Igor Mozetič
6. H2020 - HBP SGA1; Projekt „Človeški možgani“
Evropska komisija
prof. dr. Sašo Džeroski
7. H2020 - TRUE; Poti prehoda v trajnostne sisteme stročnic v Evropi
Evropska komisija
prof. dr. Marko Debeljak
8. H2020 - TomRes; Nov celosten pristop k povečanju tolerance rastlin na posamične in kombinirane strese na primeru paradiznika
Evropska komisija
prof. dr. Marko Debeljak
9. H2020 - LANDMARK; Upravljanje tal: vrednotenje, raziskave in baza znanja
Evropska komisija
prof. dr. Marko Debeljak
10. H2020 - NARSIS; Nov pristop za varnostne izboljšave reaktorjev
Evropska komisija
prof. dr. Marko Bohanec
11. H2020 - SAAM; Podpora aktivnemu staranju z večstransko pomočjo
Evropska komisija
doc. dr. Bernard Ženko
12. H2020 - PD_manager; Mobilna platforma v podporo bolnikom s parkinsonovo boleznijo
Evropska komisija
dr. Dragana Miljković
13. H2020 - HeartMan; Osebn sistem za podporo odločanju bolnikov s srčnim popuščanjem – Srčnik
Evropska komisija
prof. dr. Marko Bohanec

14. H2020 - CF-Web; CloudFlows spletno tržišče za podatkovno in tekstovno analitiko
Evropska komisija
doc. dr. Martin Žnidaršič
15. Rudarjenje kompleksnih podatkov s področja znanosti o življenju in znanosti o okolju
Javna agencija za raziskovalno dejavnost RS
prof. dr. Sašo Džeroski
16. Rudarjenje heterogenih povezanih biomedicinskih podatkov
Javna agencija za raziskovalno dejavnost RS
prof. dr. Nada Lavrač
17. Označevanje semantičnih vlog v slovenščini in hrvaščini
Javna agencija za raziskovalno dejavnost RS
prof. dr. Tomaž Erjavec

PROGRAM

1. Tehnologije znanja
prof. dr. Nada Lavrač

PROJEKTI

1. Integrativne raziskave evolucije spolnega dimorfizma
prof. dr. Sašo Džeroski
2. Viri, orodja in metode za raziskovalne nestandardne spletne slovenščine
prof. dr. Tomaž Erjavec
3. Molekulske osnove interakcij med vinsko trto in fitoplazemskimi povzročitelji trsnih rumenic
prof. dr. Nada Lavrač
4. Razvoj sodobnih metodoloških pristopov za preučevanje vedenja prostoživečih živali: raziskave nastanka problematičnih medvedov v kontrastnih okoljih Evrope
prof. dr. Sašo Džeroski
5. Vpliv formalnih in neformalnih komunikacij na trg kapitala
dr. Senja Pollak
6. Kolokacije kot temelj jezikovnega opisa: semantični in časovni vidiki
dr. Nikola Ljubešić
7. Slovenska znanstvena besedila: viri in opis
prof. dr. Tomaž Erjavec
8. Analiza heterogenih informacijskih omrežij za odkrivanje zakonitosti v znanostih o življenju
prof. dr. Nada Lavrač
9. Strojno učenje za sistemske znanosti
prof. dr. Sašo Džeroski
10. Viri, metode in orodja za razumevanje, prepoznavanje in razvrščanje različnih oblik družbeno nesprejemljivega diskurza v informacijski družbi
prof. dr. Tomaž Erjavec
11. Gostovanje pri vodji ERC projekta
dr. Vedrana Vidulin

12. Prepovedane knjige na Slovenskem v zgodnjem novem veku
prof. dr. Tomaž Erjavec
13. Napovedovanje strukturiranih vrednosti z uporabo v trajnostni kmetijski proizvodnji
prof. dr. Sašo Džeroski
14. F4F: Funkcionalna živila prihodnosti
doc. dr. Bernard Ženko
15. TRAIN: Big data in modeli bolezni: Čezmejna platforma validiranih kompletov za biotehnoško industrijo
prof. dr. Sašo Džeroski
16. Rudarjenje podatkov in podpora odločanja pri trajnostni pridelavi hrane
dr. Vladimir Kuzmanovski
17. Učenje modelov bolezni in terapij za sistemsko in osebno medicino
dr. Jovan Tanevski
18. ReLDI - Inicijativa za jezikoslovne podatke
dr. Nikola Ljubešić
19. ECML PKDD 2017 - 28. Evropska konferenca s področja strojnega učenja in 12. Evropska konferenca s področja odkrivanja znanja iz podatkov; Skopje, Makedonija, 18.-22. 09. 2017
prof. dr. Sašo Džeroski

20. ECML PKDD 2017 - 28. Evropska konferenca s področja strojnega učenja in 12. Evropska konferenca s področja odkrivanja znanja iz podatkov; Skopje, Makedonija, 18.-22. 09. 2017
prof. dr. Sašo Džeroski
21. CLARIN projekt: Razvoj in integracija večjezičnega označevanja korpusov
prof. dr. Tomaž Erjavec

VEČJA NOVA POGODBENA DELA

1. Izvedba seminarja „Data Mining in CloudFlows and TextFlows“
Comtrade, d. o. o.
prof. dr. Nada Lavrač
2. Izdelava prototipne programske rešitve za podporo pol-avtomatskemu luščenju in upravljanju terminologije enojezičnih in dvojezičnih korpusov
Iolar, d. o. o.
dr. Senja Pollak
3. Izdelava projekta „TermIOLAR-2“: Prototip programske rešitve problema luščenja in poravnave terminologije iz vzporednih korpusov pomnilnikov prevodov
Iolar, d. o. o.
dr. Senja Pollak

OBISKI

1. prof. dr. Donato Malerba, Università degli Studi di Bari "Aldo Moro", Italija, 13.-15. 3. 2017
2. prof. dr. Michelangelo Ceci, Università degli Studi di Bari "Aldo Moro", Bari, Italija 2.-8. 4. 2017 in 1.-2. 6. 2017
3. dr. Ivica Dimitrovski, Ss. Cyril and Methodius, Faculty of Computer Science and Engineering, University of Skopje, Skopje, Makedonija, 2.-8. 4. 2017
4. dr. Tomislav Šmuc, Institut Ruder Bošković, Zagreb, Hrvaška, 3.-5. 4. 2017
5. prof. dr. Hendrik Blockeel, K. U. Leuven, Department of Computer Science, Heverlee, Belgija, 31. 5.-3. 6. 2017
6. Franziska Schütze, Global Climate Forum, Berlin, Nemčija, 16.-21. 5. 2017
7. Dan Davis, SHARK company, Boston, ZDA, 5. 6. 2017
8. dr. Vesna Andova, Faculty of Electrical Engineering and Information Technologies, Skopje, Makedonija, 17.-25. 6. 2017
9. Ilin Tolovski, Faculty of Electrical Engineering and Information Technologies, Skopje, Makedonija, 17.-25. 6. 2017
10. Ana Kostovska, Faculty of Electrical Engineering and Information Technologies, Skopje, Makedonija, 17.-25. 6. 2017
11. dr. Maja Miličević, Univerza v Beogradu, Filološka fakulteta, Oddelek za splošno jezikoslovje, Beograd, Srbija, 20.-25. 6. 2017
12. dr. Tanja Samardić, University of Zurich, URPP Language and Space, Zürich, Švica, 20.-25. 6. 2017
13. dr. Saturnino Luz, Usher Institute of Population Health Sciences and Informatics, Edinburgh Medical School: Molecular, Genetic and Population Health Sciences, 25.-27. 6. 2018 ter 21.-22. 9. 2017
14. Franklin Parrales Bravo, Complutense University of Madrid, Madrid, Španija, 24. 9.-6. 12. 2017
15. prof. dr. Geoff Squire, James Hutton Institute, Dundee, Škotska, 4.-6. 9. 2017
16. Nataša Terzić, Center za razvoj zdravstvenega sistema, Institut za varovanje zdravja, Podgorica, Črna gora, 16.-19. 10. 2017
17. Jihed Khiari z NEC Europe Ltd, Heidelberg, Nemčija, 22.-27. 10. 2017
18. Jozef Misutka, Karlova Univerza v Pragi, Češka republika, 26.-27. 10. 2017
19. Frank van der Velde, University of Twente, Enschede, Centre for Telematics and Information Technology (CTIT), Enschede, Nizozemska, 8.-10. 11. 2017
20. prof. dr. Geraint Wiggins, Queen Mary University of London, London Velika Britanija, 15.-26. 11. 2017
21. dr. Maximilian Moser, Medical University of Graz, Avstrija, 20. 11. 2017
22. mag. Thomas Hassler, Human Research Institute, Gradec, Avstrija, 20. 11. 2017
23. prof. dr. Henrik Bostrom, KTH Royal Institute of Technology, Stockholm, Švedska, 23.-26. 11. 2017
24. dr. Luca Leonardo Bologna, Italian National Research Council (CNR) Università degli Studi di Palermo, Italija, 26.-30. 11. 2017
25. Luke Lucas, European Space Agency (ESA), Darmstadt, Nemčija, 10.-14. 12. 2017
26. Redouane Boumghar, European Space Agency (ESA), Darmstadt, Nemčija, 10.-14. 12. 2017

SEMINARJI IN PREDAVANJA NA IJS

1. Vid Podpečan: Predstavitel robota NAO za robotski krožek iz OŠ Mirana Jarca, 30. 3. 2017
2. Vid Podpečan: Predstavitel robota NAO na osnovnošolce in ostale prijavljene, 7. 11. 2017

UDELEŽBA NA ZNANSTVENIH ALI STROKOVNIH ZBOROVANJIH

1. Tina Anžič, Milica Bauer, Martin Breskvar, Sašo Džeroski, Dragi Kocev, Jan Kralj, Kralj Novak Petra, Vladimir Kuzmanovski, Nada Lavrač, Biljana Mileva Boshkoska, Aljaž Osojnik, Panče Panov, Matej Petković, Tomaž Stepišnik Perdih, Jovan Tanevski, Aneta Trajanov, Vedrana Vidulin, Bernard Ženko, Martin Žnidaršič, organizacija conference ter udeležba na ECML PKDD 2017 in udeležba na zaključnem sestanku EU projekta MAESTRA (projektne predstavitve), Skopje, Makedonija, 13.-23. 9. 2017
2. Milica Bauer, Panče Panov, delovni sestanek za organizacijo conference ECML/PKDD 2018, Skopje, Makedonija, 17.-20. 4. 2017
3. Milica Bauer, ADMA 2017, Bernardin, Slovenija, 18.-20. 5. 2017
4. Marko Bohanec: IWoMCDM'17, International Workshop on Multiple Criteria Decision Making'17, Ustroń, Poland, 2.-4. 4. 2017 (vabljeni predavatelj)
5. Marko Bohanec: ICDSST 2017, EWG-DSS 2017 International Conference on Decision Support System Technology, Namur, Belgija, 29.-31. 5. 2017 (poster)
6. Marko Bohanec: KM 2017, Knowledge Management 2017, Novo mesto, Slovenija, 21.-24. 6. 2017 (vabljeni predavatelj)
7. Marko Bohanec: MCDM 2017, 24th International Conference on Multiple Criteria Decision Making, Ottawa, Kanada, 10.-14. 7. 2017 (referat)
8. Marko Bohanec: SOR 2017, The 14th International Symposium on Operations Research in Slovenia, Bled, Slovenija, 27.-29. 9. 2017 (referat, soorganizator sekcije)
9. Marko Bohanec, Vladimir Kuzmanovski, Biljana Mileva Boshkoska, Matej Petković, Senja Pollak, Jovan Tanevski, Aneta Trajanov: IS 2017, Informacijska družba 2017, Slovenian Conference on Artificial Intelligence, Ljubljana, Slovenija, 12.-13. 10. 2017 (6 referatov)
10. Marko Bohanec, Fakulteta organizacije in informatike, Varaždin, Hrvaška, 5. 12. 2017 (vabljeni predavatelj)
11. Martin Breskvar, Igor Mozetič, Intl. Conf. on Data Science and Advanced Analytics, Tokyo, Japan, 19.-21. 2. 2017
12. Marko Debeljak, Konferenca Landscape management for functional biodiversity, Dundee, Velika Britanija, 28. 3.-1. 4. 2017
13. Marko Debeljak, Aneta Trajanov, LANDMARK delovni sestanek, Wageningen, Nizozemska, 7.-10. 3. 2017 (pasivno)
14. Marko Debeljak, Aneta Trajanov, TRUE kick-off sestanek, Edinburgh, Škotska, 18.-22. 4. 2017 (pasivno)
15. Marko Debeljak, Aneta Trajanov, TomRes kick-off setanek, Torino, Italija, 7.-10. 6. 2017 (pasivno)
16. Marko Debeljak, Vladimir Kuzmanovski, Aneta Trajanov, BIODIV delovni sestanek, Pariz Francija, 19.-21. 6. 2017 (pasivno)
17. Marko Debeljak, Vladimir Kuzmanovski, Aneta Trajanov, LANDMARK delovni sestanek, Parma, Italija, 22.-25. 10. 2017 (pasivno)
18. Marko Debeljak, Aneta Trajanov, TRUE delovni sestanek, Hohenheim, Nemčija, 20.-22. 11. 2017 (pasivno)
19. Marko Debeljak, Aneta Trajanov, BES konferenca, Ghent, Belgija, 10.-14. 12. 2017 (članek)
20. Sašo Džeroski, Pattern Recognition and Computer Vision Colloquium, Praga, Češka, 18.-19. 1. 2017 (vabljeni predavatelj)
21. Sašo Džeroski, delavnica HBP SGA2 planning workshop to align the SGA2 plans of the SPs, Malaga, Španija, 19.-22. 2. 2017
22. Sašo Džeroski, sestanek HBP Stakeholder Board in obisk projektnih partnerjev, Ženeva, Lozana, Švica, 29.-31. 3. 2017 (pasivno)
23. Sašo Džeroski, sestanek EU-projekta MAESTRA, Skopje, Makedonija, 17.-18. 4. 2017 (projektne predstavitve)

24. Sašo Džeroski, delavnica HORIZON 2020 Writing Successful ERC Advanced Grant Proposals, Amsterdam, Nizozemska, 27.–28. 4. 2017 (pasivno)
25. Marko Debeljak, Sašo Džeroski, Aneta Trajanov, LANDMARK Roundtable 5 (projektna predstavitev) in obisk NEC Laboratories Europe (vabljeni predavanja), Cluj-Napoca in Heidelberg, Romunija, Nemčija, 30. 5.–1. 6. 2017
26. Sašo Džeroski, recenzentski sestanek EU-projekta LANDMARK (projektna predstavitev) in delovni sestanek na KUL, Bruselj, Leuven, Belgija, 7.–9. 6. 2017
27. Darko Aleksovski, Tina Anžič, Milica Bauer, Marko Bohanec, Martin Breskvar, Sašo Džeroski, Tomaž Erjavec, Petra Kralj Novak, Dragi Kocev, Jan Kralj, Vladimir Kuzmanovski, Nada Lavrač, Jurica Levatič, Matej Martinc, Biljana Mileva Boshkoska, Dragana Miljković, Aljaž Osojnik, Nikola Simidjievski, Tomaž Stepišnik Perdiš, Panče Panov, Matej Petković, Vid Podpečan, Senja Pollak, Jasmina Smailović, Anže Vavpetič, Vedrana Vidulin, Jovan Tanevski, Katja Zupan, Martin Žnidaršič, Bernard Ženko, udeležba na 10th International Ljubljana-Zagreb Workshop on Knowledge Technologies, Fiesca, Slovenija, 13.–14. 6. 2017 (predstavitev)
28. Sašo Džeroski, 4th HBP School – Future Computing: Brain Science and Artificial Intelligence, Obergurgl, Avstrija, 14.–17. 6. 2017 (2 vabljeni predavanja)
29. Sašo Džeroski, Letna konferenca Evropske Akademije znanosti, Budimpešta, Madžarska, 4.–6. 9. 2017 (pasivno)
30. Sašo Džeroski, Panče Panov, uvodni sestanek INTERREG V-A 2014–2020 projekta TRAIN, Trst, Italija, 26. 9. 2017 (projektna predstavitev)
31. Sašo Džeroski, sestanek EU-projekta COSMOS, Brioni, Hrvaška, 8.–10. 10. 2017 (vabljeni predavanja)
32. Martin Breskvar, Sašo Džeroski, Igor Mozetič, Matej Petković, Jovan Tanevski, Tomaž Stepišnik Perdiš, Discovery Science 2017 (4 referati), Data Science and Advanced Analytics 2017 (pasivno), Kyoto, Japonska, 15.–21. 10. 2017
33. Sašo Džeroski, Intelligent Data Analysis 2017, London, Velika Britanija, 25.–29. 10. 2017 (referat)
34. Sašo Džeroski, delavnica Data Science and Advanced Analytics 2017, Zagreb, Hrvaška, 29.–30. 11. 2017 (vabljeni predavanja)
35. Tomaž Erjavec, sestanek projekta Clarin, Gothenburg, Švedska, 20.–22. 6. 2017
36. Tomaž Erjavec, Nikola Ljubešić, ACL 2017, Vancouver, Kanada, 30. 7.–11. 8. 2017 (referat)
37. Tomaž Erjavec, Biographical data in a Digital World, Linz, Avstrija, 5.–7. 11. 2017 (referat)
38. Tomaž Erjavec, Clarin General Assembly meeting, Helsinki, Finska, 27.–30. 11. 2017
39. Petra Kralj Novak, sestanek projekta Dolfin, Zürich, Švica, 10.–12. 1. 2017
40. Petra Kralj Novak, recenzijski sestanek projekta Simpol, Bruselj, Belgija, 25.–26. 4. 2017
41. Petra Kralj Novak, Big Data: New challenges for law and ethics, Ljubljana, 22.–23. 5. 2017 (pasivno)
42. Jan Kralj, Petra Kralj Novak, Igor Mozetič, Complex Networks 2017, Lyon, Francija, 28. 11.–2. 12. 2017 (3 referati)
43. Vladimir Kuzmanovski, delavnica Open Land Data, Wageningen, Amsterdam, 29. 6.–4. 7. 2017
44. Vladimir Kuzmanovski, ICEI 2017, Helsinki, Finska, 1.–7. 8. 2017 (1 referat)
45. Nada Lavrač, ISWC 2017, Dunaj, Avstrija, 20.–25. 10. 2017 (vabljeni predavanja)
46. Matej Martinc, Nada Lavrač, YSU SSS 4th International Conference, Yerevan, Armenija, 30. 9.–8. 10. 2017 (referat)
47. Matej Martinc, ESWC 2017, Portorož, Slovenija, 28.–29. 5. 2017
48. Matej Martinc, ICCS 2017, Atlanta, ZDA, 18.–25. 6. 2017 (1 referat)
49. Matej Martinc, CLEF 2017, Dublin, Irsko, 10.–15. 9. 2017 (referat)
50. Biljana Mileva Boshkoska, European working group on decision, support systems, Namur, Belgija, 29.–31. 5. 2017 (referat)
51. Biljana Mileva Boshkoska, Slovenska konferenca o umetni inteligenci, Ljubljana, Slovenija, 9.–13. 10. 2017 (referat)
52. Biljana Mileva Boshkoska, 9th International Conference on Information Technologies and Information Society ITIS2017, 8.–10. 11. 2017, Novo mesto, Slovenija (4 referati)
53. Biljana Mileva Boshkoska, sestanek projekta PD_Manager, 20.–22. 1. 2017, Ioannina, Grčija
54. Igor Mozetič, Intl. Conf. on Synthetic Populations, Lucca, Italija, 22.–23. 2. 2017 (2)
55. Panče Panov, EUDAT Summer school on Data Management, Heraklion Kreta, Grčija
56. Matej Petković, Nikola Simidjievski, SMC-IT 2017, Alcalá de Henares, Španija, 26. 9.–1. 10. 2017 (referat)
57. Vid Podpečan, predstavitev robota NAO in odseka E8 na celjskem sejmu, Celje, Slovenija, 12.–14. 9. 2017
58. Vid Podpečan, predstavitev robota NAO na sejmu visoke tehnologije v sklopu Festivala znanosti Nova Gorica, Nova Gorica, Slovenija, 6. 10. 2017
59. Senja Pollak, International Corpus Linguistics Conference, 24.–28. 7. 2017, Birmingham, UK (1)
60. Senja Pollak, International Corpus Linguistics Summer School, 17.–21. 7. 2017, Birmingham, Velika Britanija (referat)
61. Senja Pollak, Jasmina Smailović, Martin Žnidaršič, Discourse approaches to financial communication, 3.–6. 7. 2017, Lugano, Švica (referat)
62. Senja Pollak, Jasmina Smailović, 8th International conference on computational linguistics and intelligent text processing (CICLing), 17.–23. 2017, Budapest, Hungary (2 posterja)
63. Nada Lavrač, Aljaž Osojnik, Senja Pollak, Jasmina Smailović, Bernard Ženko, Martin Žnidaršič, Kick off sestanek projekta SAAM, 11.–14. 10. 2017 Sofija, Bolgarija
64. Tomaž Stepišnik Perdiš, 3rd International Winter School on Big Data, Bari, Italija, 13. 2.–17. 2. (pasivno)
65. Tomaž Stepišnik Perdiš, 2nd Summer School on Data Science, Split, Hrvaška, 25.–29. 9. 2017 (pasivno)
66. Dragi Kocev, Nikola Simidjievski, Jovan Tanevski HBP Student Conference, Dunaj, Avstrija, 8.–10. 2. 2017 (pasivno)
67. Jovan Tanevski, Systems Medicine Conference in Slovenia: National Awareness Event, 12th CFGBC Symposium, and "Systems Medicine" Workshop, Ljubljana, 8.–9. 6. 2017 (pasivno)
68. Nikola Simidjievski, Jovan Tanevski delavnica Understanding the Brain: Neurobiology for Non-Specialists, Innsbruck, Avstrija, 3.–5. 7. 2017
69. Nikola Simidjievski, Bernard Ženko, 5th Annual HBP Summit, Glasgow, Velika Britanija, 16.–22. 10. 2017
70. Jovan Tanevski delavnica New Horizons in Clinical Neuroscience: Brain Medicine for Non-Specialists, Innsbruck, Avstrija, 5.–7. 7. 2017 (pasivno)
71. Jovan Tanevski, Aneta Trajanov, Informacijska družba, Ljubljana, 9.–13. 10. 2017 (3 referati)
72. Aneta Trajanov, delovni sestanek projekta ARENA, La Roche sur Yon, Francija, 14.–15. 12. 2017 (pasivno)
73. Anita Valmarska, AIME 2017, Dunaj, Avstrija, 20.–24. 6. 2017 (referat)
74. Matej Martinc, Anita Valmarska, poletna šola deep Learning 2017, Bilbao, Španija, 16.–22. 7. 2017
75. Katja Zupan, Seminar ReLDI "Empirični podatki v jezikoslovju: Od zasnove raziskave do analize podatkov", Ljubljana, 21.–23. 6. 2017 (pasivno)
76. Bernard Ženko, seminar projekta HBP, Pariz, Francija, 10.–12. 12. 2017

RAZISKOVALNO DELO V TUJINI

1. Darko Aleksovski, Marko Bohanec, Biljana Mileva Boshkoska, Dragana Miljković, Anita Valmarska: delovni sestanek pri partnerju PD manager projekta, Ioannina, Grčija, 19.–22. 1. 2017 (raziskovalno delo)
2. Marko Bohanec, Tehnična Univerza Poznan, Poznan, Poljska, 15.–20. 10. 2017 (raziskovalno delo)
3. Marko Bohanec, Vid Podpečan: Munchen, Nemčija, 27.–28. 11. 2017 (projektni sestanek projekta PD_manager)
4. Tomaž Erjavec, Fakulteta za humanistiko in družboslovje, Univerza v Zagrebu, 29. 9.–2. 11. 2017 (raziskovalno delo)
5. Biljana Mileva Boshkoska: University of Plymouth, UK, 27. 11.–31. 12. 2017 (raziskovalno delo)
6. Nada Lavrač, Ghent University, Ghent, Belgija, 23.–24. 2. 2017 (raziskovalno delo)
7. Nada Lavrač, Univerza v Trstu, Italija, 28.–30. 6. 2017 (priprava projekta)
8. Nikola Ljubešić, Univerza v Groningenu ter Univerza v Amsterdamu, 22.–30. 6. 2017 (raziskovalno delo)
9. Nikola Ljubešić, University of Groningen, Nizozemska, 13.–18. 11. 2017 (raziskovalno delo)
10. Vid Podpečan: Expectrum, Vasteras, Švedska, 23.–27. 10. 2017 (programiranje in animiranje NAO robota za projekt "The ROBOT" v okviru projekta INNODAY)

SODELAVCI

Raziskovalci

1. prof. dr. Marko Bohanec, znanstveni svetnik - strokovni sekretar odseka
2. prof. dr. Bojan Cestnik*
3. prof. dr. Marko Debeljak
4. prof. dr. Sašo Džeroski, znanstveni svetnik
5. prof. dr. Tomaž Erjavec, znanstveni svetnik
6. dr. Darja Fišer*
7. dr. Dragi Kocev
8. **prof. dr. Nada Lavrač, znanstveni svetnik - vodja odseka**
9. prof. dr. Zoran Levnajič*
10. dr. Nikola Ljubešić
11. prof. dr. Igor Mozetič

12. prof. dr. Ljupčo Todorovski*, znanstveni svetnik - vodja raziskovalne skupine
13. prof. dr. Tanja Urbančič*, znanstveni svetnik
14. prof. dr. Špela Vintar*, znanstveni svetnik
15. doc. dr. Bernard Ženko
16. doc. dr. Martin Žnidaršič, strokovni sekretar odseka

Podoktorski sodelavci

17. dr. Darko Aleksovski, odšel 1. 9. 2017
18. dr. Miha Grčar, odšel 1. 12. 2017
19. dr. Petra Kralj Novak
20. dr. Janez Kranjc
21. dr. Vladimir Kuzmanovski
22. dr. Jurica Levatič, odšel 1. 8. 2017
23. doc. dr. Biljana Mileva Boshkoska
24. dr. Dragana Miljković
25. dr. Blaž Mramor

26. doc. dr. Panče Panov
27. dr. Vid Podpečan
28. dr. Senja Pollak
29. dr. Nikola Simidjievski
30. dr. Jasmina Smailović
31. dr. Jovan Tanevski
32. dr. Aneta Trajanov
33. dr. Anže Vavpetič
34. dr. Vedrana Vidulin*

Mlajši raziskovalci

35. Martin Breskvar, univ. dipl. inž. rač. in inf.
36. dr. Jan Kralj
37. Matej Martinc*, univ. dipl. fil. in soc. kult.
38. Aljaž Osojnik, mag. mat.
39. Matej Petković, mag. mat.
40. Tomaž Stepišnik Perdiš, mag. mat.
41. Tadej Škvorc, mag. inž. rač. in inf.
42. Anita Valmarska, univ. dipl. inž. rač. in inf.
43. Katja Zupan, prof. angl. in slov.

Strokovni sodelavci

44. Milica Bauer, dipl. ekon.

Tehniški in administrativni sodelavci

45. Tina Anžič, Bolonjski študij II. stopnja

Opomba

* delna zaposlitev na IJS

SODELUJOČE ORGANIZACIJE

1. Aalto University, Aalto, Finska
2. ARVALIS - Institut du végétal, Pariz, Francija
3. Autonomous University of Barcelona, Španija
4. BOERSE STUTTGART HOLDING GMBH, Stuttgart, Nemčija
5. Czech Technical University, Praga, Češka Republika
6. Ecole Polytechnique Federale de Lausanne, Switzerland, Laboratory of Movement Analysis and Measurement, Lausanne, Švica
7. ESEIA, Graz Office, Mandellstraße 11/II, A-8010 Graz, Avstrija
8. ETH Eidgenössische Technische Hochschule Zürich, Zürich, Švica
9. Fakulteta za informacijske študije, Novo mesto
10. Fundació Privada Barcelona Digital Centre Tecnològic, BDIGITAL, Barcelona, Španija
11. Global Climate Forum, Nemčija
12. Goldsmiths College, University of London, London, Velika Britanija
13. Gozdarski inštitut Slovenije, Ljubljana
14. Helsinki Institute for Information Technology, Helsinki, Finska
15. Honda R&D Europe (Deutschland) GmbH, Offenbach, Nemčija
16. INESC TEC, Porto, Portugalska
17. INRA Eco-INNOV, Grignon, Francija
18. Institut Rudjer Bošković, Zagreb, Hrvaška
19. Institute for Advanced Studies, IMT Lucca, Italija

20. Katholieke Universiteit Leuven, Department of Computer Science, Leuven, Belgija
21. Klinik für Neurologie, UniversitätsSpital Zürich, Switzerland, Clinical Neurorehabilitation, Zürich, Švica
22. Kmetijski inštitut Slovenije, Ljubljana
23. London Institute for Mathematical Sciences, LIMS, Velika Britanija
24. Mednarodna podiplomska šola Jožefa Stefana
25. Nacionalni inštitut za biologijo, Ljubljana
26. OLAI, d. o. o., Pot za Brdom 100, SI-1000 Ljubljana, Slovenija
27. Queen Mary University of London, Velika Britanija
28. Sowa Labs GmbH, Stuttgart, Nemčija
29. Temida, d. o. o., Ljubljana
30. The Institute of Scientific and Industrial Research Osaka University, Osaka, Japonska
31. The National Research Council, Rim, Italija
32. Universidade Nova da Lisboa, Lisboa, Portugalska
33. Università degli Studi di Milano, Italy, Applied Intelligent Systems Laboratory Department of Computer Science, Milano, Italija
34. Università degli Studi di Padova, Italy, Department of General Psychology, Padova, Italija
35. Universitaet, Zürich, Švica
36. Université Paris, Pantheon-Sorbonne, Francija
37. University College, Dublin, Irsko
38. University of Bari, Department of informatics, Bari, Italija
39. University of Cambridge, Cambridge, Velika Britanija
40. University of Coimbra, Coimbra, Portugalska
41. University of Helsinki, Helsinki, Finska
42. University of Oxford, Oxford, Velika Britanija
43. University of York, Department of Computer Science, York, Velika Britanija
44. University of New South Wales, School of Computer Science and Engineering, Avstralija
45. Université Paris, Pantheon-Sorbonne, Pariz, Francija
46. University of Porto, Artificial Intelligence and Computer Science Laboratory, Portugalska
47. University of Skopje, Faculty of Computer Science and Engineering, Makedonija
48. University of Twente, Enschede, Nizozemska
49. Univerza v Beogradu, Filološki fakultet, Beograd, Srbija
50. Univerza v Novi Gorici
51. Univerza v Ljubljani, Biotehniška fakulteta, Ljubljana
52. Univerza v Ljubljani, Fakulteta za družbene vede, Ljubljana
53. Univerza v Ljubljani, Filozofska fakulteta, Ljubljana
54. Univerza v Ljubljani, Fakulteta za matematiko in fiziko, Ljubljana
55. Univerza v Ljubljani, Fakulteta za računalništvo in informatiko, Ljubljana
56. Univerza v Ljubljani, Fakulteta za gradbeništvo in geodezijo, Ljubljana
57. Univerza v Ljubljani, Fakulteta za upravo, Ljubljana
58. Univerza v Mariboru, Fakulteta za elektrotehniko, računalništvo in informatiko, Maribor
59. Univerza v Mariboru, Fakulteta za kmetijstvo in biosistemske vede, Pivola
60. Univerza v Mariboru, Medicinska fakulteta, Maribor
61. Univerza v Zagrebu, Fakulteta za humanistiko in družboslovje, Zagreb, Hrvaška
62. Univerza v Ženevi, Center za računalništvo, Ženeva, Švica
63. Zavod za zdravstveno varstvo Maribor
64. ZRC SAZU, Inštitut za slovensko literaturo in literarne vede
65. XLAB, d. o. o., Ljubljana

BIBLIOGRAFIJA

IZVIRNI ZNANSTVENI ČLANEK

1. Kristina Ban, Matjaž Perc, Zoran Levnajić, "Robust clustering of languages across Wikipedia growth", *Royal Society Open Science*, **4**, 10, 1-12, 2017. [COBISS.SI-ID 23440904]
2. Marko Bohanec, Nejc Trdin, Branko Kontić, "A qualitative multi-criteria modelling approach to the assessment of electric energy production technologies in Slovenia", *Cent. Eur. j. oper. res.*, **25**, 3, 611-625, 2017. [COBISS.SI-ID 29867815]
3. Bojan Cestnik, Elsa Fabbretti, Donatella Gubiani, Tanja Urbančič, Nada Lavrač, "Reducing the search space in literature-based discovery by exploring outlier documents: a case study in finding links between gut microbiome and Alzheimer's disease", *Genomics and computational biology*, **3**, 3, e58, 2017. [COBISS.SI-ID 30497575]
4. Dagmar Divjak, Serge Sharoff, Tomaž Erjavec, "Slavic corpus and computational linguistics", *J. Slav. linguist.*, **25**, 2, 171-198, 2017. [COBISS.SI-ID 31281191]
5. Peter Gabrovšek, Darko Aleksovski, Igor Mozetič, Miha Grčar, "Twitter sentiment around the earnings announcement events", *PloS one*, **12**, 2, e0173151, 2017. [COBISS.SI-ID 30277671]
6. Dragan Gamberger *et al.* (5 avtorjev), "Identification of clusters of rapid and slow decliners among subjects at risk for Alzheimer's disease", *Sci. rep.*, **7**, 6763, 2017. [COBISS.SI-ID 4875771]
7. Miha Grčar, Darko Čerepnalkoski, Igor Mozetič, Petra Kralj Novak, "Stance and influence of Twitter users regarding the Brexit referendum", *Comput. soc. networks*, **4**, 6, 2017. [COBISS.SI-ID 30650663]
8. Donatella Gubiani, Elsa Fabbretti, Bojan Cestnik, Nada Lavrač, Tanja Urbančič, "Outlier based literature exploration for cross-domain linking of Alzheimer's disease and gut microbiota", *Expert syst. appl.*, **85**, 386-396, 2017. [COBISS.SI-ID 4787451]
9. Sabina Horvat, Aleksander Mahnič, Martin Breskvar, Sašo Džeroski, Maja Rupnik, "Evaluating the effect of Clostridium difficile conditioned medium on fecal microbiota community structure", *Sci. rep.*, **7**, 16448, 2017. [COBISS.SI-ID 512751416]
10. Jernej Jevšenak, Sašo Džeroski, Tom Levanič, "Uporaba metod strojnega učenja za preučevanje odnosov med značilnostmi branik in okoljem", *Acta silvae et ligni*, [Št.] 114, 21-29, dec. 2017. [COBISS.SI-ID 4998310]
11. Manja Klemenčič, Martin Žnidaršič, Anže Vavpetič, Matej Martinc, "Erasmus students' involvement in quality enhancement of Erasmus+ mobility through digital ethnography and ErasmusShouts", *Stud. high. educ. (Dorchester-on-Thames)*, **42**, 5, 925-932, 2017. [COBISS.SI-ID 30471975]
12. Janez Kranjc, Roman Orač, Vid Podpečan, Nada Lavrač, Marko Robnik Šikonja, "ClowdFlows: online workflows for distributed big data mining", *FGCS, Future gener. comput. syst.*, **68**, 38-58, 2017. [COBISS.SI-ID 29851943]
13. Marc Grau Leguia, Ralph G Andrzejak, Zoran Levnajić, "Evolutionary optimization of network reconstruction from derivative-variable correlations", *J. phys. A, Math. theor.*, **50**, 33, 334001, 2017. [COBISS.SI-ID 2048462099]
14. Jurica Levatić, Michelangelo Ceci, Dragi Kocev, Sašo Džeroski, "Self-training for multi-target regression with tree ensembles", *Knowl.-based syst.*, **123**, 41-60, 2017. [COBISS.SI-ID 30498855]
15. Jurica Levatić, Michelangelo Ceci, Dragi Kocev, Sašo Džeroski, "Semi-supervised classification trees", *Journal of intelligent information systems*, **49**, 3, 461-486, 2017. [COBISS.SI-ID 30539559]
16. Matej Mihelčič, Goran Šimić, Mirjana Babić Leko, Nada Lavrač, Sašo Džeroski, Tomislav Šmuc, "Using redescription mining to relate clinical and biological characteristics of cognitively impaired and Alzheimer's disease patients", *PloS one*, **12**, 10, 0187364, 2017. [COBISS.SI-ID 30880295]
17. Matej Mihelčič, Marko Bohanec, "Approximating incompletely defined utility functions of qualitative multi-criteria modeling method DEX", *Cent. Eur. j. oper. res.*, **25**, 3, 627-649, 2017. [COBISS.SI-ID 29661735]
18. Matej Mihelčič, Sašo Džeroski, Nada Lavrač, Tomislav Šmuc, "A framework for redescription set construction", *Expert syst. appl.*, **68**, 196-215, 2017. [COBISS.SI-ID 29905191]
19. Dragana Miljković, Nada Lavrač, Marko Bohanec, Biljana Mileva-Boshkoska, "Discovering dependencies between domains of redox potential and plant defence through triplet extraction and copulas", *Int. j. intell. eng. inform.*, 1-17, 2017. [COBISS.SI-ID 2048463379]
20. Jasmina Smailović, Martin Žnidaršič, Aljoša Valentinčič, Igor Lončarski, Marko Pahor, Pedro Martins, Senja Pollak, "Automatic analysis of annual financial reports: a case study", *Computación y sistemas*, **21**, 4, 809-818, 2017. [COBISS.SI-ID 24345318]
21. Erik Tjong Kim Sang *et al.* (19 avtorjev), "The CLIN27 shared Task: Translating historical text to contemporary language for improving automatic linguistic annotation", *Computational linguistics in the Netherlands journal*, **7**, 53-64, 2017. [COBISS.SI-ID 31270951]
22. Antonio Toral, Miquel Esplá-Gomis, Filip Klubička, Nikola Ljubešić, Vassilis Papavassiliou, Prokopis Prokopidis, Raphael Rubino, Andy Way, "Crawl and crowd to bring machine translation to under-resourced languages", *Language resources and evaluation*, **51**, 4, 1019-1051, 2017. [COBISS.SI-ID 30359079]
23. Kostas M. Tsiouris *et al.* (11 avtorjev), "PD_Manager: a mHealth platform for Parkinson's disease patient management", *Healthc. technol. lett.*, **4**, 3, 102-108, 2017. [COBISS.SI-ID 30490151]
24. Tea Tušar, Klemen Gantar, Valentin Koblar, Bernard Ženko, Bogdan Filipič, "A study of overfitting in optimization of a manufacturing quality control procedure", *Applied soft computing*, **59**, 77-87, 2017. [COBISS.SI-ID 30552359]
25. Anita Valmarska, Nada Lavrač, Johannes Fürnkranz, Marko Robnik Šikonja, "Refinement and selection heuristics in subgroup discovery and classification rule learning", *Expert syst. appl.*, **81**, 147-162, 2017. [COBISS.SI-ID 30400807]

OBJAVLJENI ZNANSTVENI PRISPEVEK NA KONFERENCI

1. Michael Beißwenger *et al.* (11 avtorjev), "Closing a gap in the language resources landscape: groundwork and best practices from projects on computer-mediated communication in four European countries", V: *Selected papers from the CLARIN Annual Conference 2016*, (Linköping Electronic Conference Proceedings, 136), [S. l., s. n.], 2017, 1-18. [COBISS.SI-ID 64614754]
2. Michael Beißwenger *et al.* (19 avtorjev), "Connecting resources: which issues have to be solved to integrate CMC corpora from heterogeneous sources and for different languages?", V: *Proceedings of the 5th Conference on CMC and Social Media Corpora for the Humanities (cmccorpora17)*, Egon W. Stemle, ur., Ciara R. Wigham, ur., Bolzano, [S. n.], 2017, 10-13. [COBISS.SI-ID 65428834]
3. Mate Beštek, Dragi Kocev, Sašo Džeroski, Andrej Brodnik, Rade Iljaž, "Modelling time-series of glucose measurements from diabetes patients using predictive clustering trees", V: *Artificial intelligence in medicine: 16th Conference on Artificial Intelligence in Medicine, AIME 2017, Vienna, Austria, June 21-24, 2017: proceedings*, (Lecture notes in artificial intelligence, **10259**), Annette ten Teije, ur., 2017, 95-104. [COBISS.SI-ID 30599975]
4. Marko Bohanec, "Multi-criteria DEX model: an overview and analysis", V: *SOR '17 proceedings*, Lidija Zadnik Stirn, ur., et al, Ljubljana, Slovenian Society Informatika, Section for Operational Research, 2017, 155-160. [COBISS.SI-ID 30808871]
5. Marko Bohanec, Erik Dovgan, Pavel Maslov, Aljoša Vodopija, Mitja Luštrek, Paolo Emilio Puddu, Michele Schiariti, Maria Constanza Ciancarelli, Anneleen Baert, Sofie Pardaens, Els Clays, "Designing a personal decision support system for congestive heart failure management", V: *Slovenska konferenca o umetni inteligenci: zbornik 20. mednarodne multikonference Informacijska družba - IS 2017*, 9.-13. oktober 2017, [Ljubljana, Slovenija]: zvezek A: *proceedings of the 20th International Multiconference Information Society - IS 2017*, 9th-13th October, 2017, Ljubljana, Slovenia: volume A, Mitja Luštrek, ur., Rok Piltaver, ur., Matjaž Gams, ur., Ljubljana, Institut Jožef Stefan, 2017, 67-70. [COBISS.SI-ID 30857511]
6. Martin Breskvar, Dragi Kocev, Sašo Džeroski, "Multi-label classification using random label subset selections", V: *Discovery science: 20th International Conference, DS 2017, Kyoto, Japan, October 15-17, 2017: proceedings*, (Lecture notes in artificial intelligence, **10558**), Akihiro Yamamoto, ur., 2017, 108-115. [COBISS.SI-ID 30862375]
7. Martin Breskvar, Dragi Kocev, Jurica Levatić, Aljaž Osojnik, Matej Petkovič, Nikola Simidjievski, Bernard Ženko, Redouane Boumghar, Luke Lucas, "Predicting thermal power consumption of the Mars

- Express satellite with machine learning", V: *Proceedings*, 6th IEEE International Conference on Space Mission Challenges for Information Technology, SMC-IT 2017, 27-29 September 2017, Alcalá de Henares, Spain, Danvers, IEEE = The Institute of Electrical and Electronics Engineers, 2017, 88-93. [COBISS.SI-ID 31053863]
8. Vitor Cerqueira, Luís Torgo, Jasmina Smalović, Igor Mozetič, "A comparative study of performance estimation methods for time series forecasting", V: *IEEE/ACM/ASA DSAA' 2017*, Eric Gaussier, ur., Danvers, IEEE = Institute of Electrical and Electronics Engineers, 2017, 529-538. [COBISS.SI-ID 30872103]
 9. Marko Debeljak, Vladimir Kuzmanovski, Veronique Tosser, Aneta Trajanov, "Knowledge discovery from complex ecological data: exploring Syrphidae species in agricultural landscapes", V: *Slovenska konferenca o umetni inteligenci: zbornik 20. mednarodne multikonference Informacijska družba - IS 2017*, 9.-13. oktober 2017, [Ljubljana, Slovenia]: zvezek A: *proceedings of the 20th International Multiconference Information Society - IS 2017*, 9th-13th October, 2017, Ljubljana, Slovenia: volume A, Mitja Luštrek, ur., Rok Piltaver, ur., Matjaž Gams, ur., Ljubljana, Institut Jožef Stefan, 2017, 55-58. [COBISS.SI-ID 30856743]
 10. Marko Debeljak, Florence Leprince, Sašo Džeroski, Aneta Trajanov, "Criteria for co-existence of GM and conventional Maize production", V: *Slovenska konferenca o umetni inteligenci: zbornik 20. mednarodne multikonference Informacijska družba - IS 2017*, 9.-13. oktober 2017, [Ljubljana, Slovenia]: zvezek A: *proceedings of the 20th International Multiconference Information Society - IS 2017*, 9th-13th October, 2017, Ljubljana, Slovenia: volume A, Mitja Luštrek, ur., Rok Piltaver, ur., Matjaž Gams, ur., Ljubljana, Institut Jožef Stefan, 2017, 51-54. [COBISS.SI-ID 30856487]
 11. Ivica Dimitrovski, Dragi Kocev, Suzana Loskovska, Sašo Džeroski, "Image representation annotation and retrieval with predictive clustering trees", V: *Machine learning and knowledge discovery in databases: European Conference, ECML PKDD 2017 Skopje, Macedonia, September 18-22, 2017. Part III*, (Lecture notes in artificial intelligence, **10536**), Yasemin Altun, ur., Sašo Džeroski, ur., 2017, 363-367. [COBISS.SI-ID 31157799]
 12. Kaja Dobrovoljc, Tomaž Erjavec, Simon Krek, "The Universal Dependencies Treebank for Slovenian", V: *Proceedings of the EACL workshop*, The 6th Workshop on Balto-Slavic Natural Language Processing, April 4, 2017 Valencia, Spain, Stroudsburg, The Association for Computational Linguistics = ACL, 2017, 33-38. [COBISS.SI-ID 66146402]
 13. Tomaž Erjavec, Nikola Ljubešić, Darja Fišer, "Adapting a state-of-the-art tagger for South Slavic languages to non-standard text", V: *Proceedings of the EACL workshop*, The 6th Workshop on Balto-Slavic Natural Language Processing, April 4, 2017 Valencia, Spain, Stroudsburg, The Association for Computational Linguistics = ACL, 2017, 60-68. [COBISS.SI-ID 64001634]
 14. Darja Fišer, Jakob Lenardič, "Parliamentary Corpora in the CLARIN infrastructure", V: *Proceedings*, [S. l., s. n.], 2017, 1-4. [COBISS.SI-ID 65322082]
 15. Darja Fišer, Nikola Ljubešić, Tomaž Erjavec, "Legal framework, dataset and annotation schema for socially unacceptable online discourse practices in Slovene", V: *Proceedings of the 1st Workshop on Abusive Language Online*, 1st Workshop on Abusive Language Online, ACL 2017, Vancouver, Canada, August 4 2017, Vancouver, ACL, 2017, 46-51. [COBISS.SI-ID 65015138]
 16. Valentin Gjorgjioski, Dragi Kocev, Andrej Bončina, Sašo Džeroski, Marko Debeljak, "Predictive clustering of multi-dimensional time series applied to forest growing stock data for different tree sizes", V: *Data-driven innovation: 9th International Conference, ICT Innovations 2017 Skopje, Macedonia, September 18-23, 2017: proceedings*, (Communications in computer and information science, **758**), Dimitar Trajanov, ur., Veriva Bakeva, ur., 2017, 186-195. [COBISS.SI-ID 30863143]
 17. Taja Kuzman, Darja Fišer, "Corpus-based analysis of demonyms in Slovene Twitter", V: *Proceedings of the 5th Conference on CMC and Social Media Corpora for the Humanities (cmccorpora17)*, Egon W. Stemle, ur., Ciara R. Wigham, ur., Bolzano, [S. n.], 2017, 29-33. [COBISS.SI-ID 65216098]
 18. Jurica Levatič, Marija Brbič, Tomaž Stepišnik Perdih, Dragi Kocev, Vedrana Vidulin, Tomislav Šmuc, Fran Supek, Sašo Džeroski, "Phenotype prediction with semi-supervised learning", V: *New frontiers in mining complex patterns: Sixth edition of the International Workshop NFMCP 2017 in conjunction with ECML-PKDD 2017, September 18-22, 2017, Skopje, Macedonia*, 2017. [COBISS.SI-ID 30862119]
 19. Nikola Ljubešić, Tomaž Erjavec, Darja Fišer, Enhard Hinrichs, Marie Hinrichs, Cyprian Laskowski, Filip Petkovski, Wei Qui, "Multilingual text annotation of Slovenian, Croatian and Serbian with WebLicht", V: *Proceedings*, CLARIN Annual Conference 2017, 18-20 September, Budapest, Hungary, 2017, 18. [COBISS.SI-ID 31083303]
 20. Nikola Ljubešić, Darja Fišer, Tomaž Erjavec, "Language-independent gender prediction on Twitter", V: *Proceedings of the Second Workshop on Natural Language Processing and Computational Social Science*, The 2nd Workshop on Natural Language Processing and Computational Social Science, August 4, Vancouver, Canada, Vancouver, Association for Computational Linguistics, 2017, 1-6. [COBISS.SI-ID 65025378]
 21. Žiga Lukšič, Jovan Tanevski, Sašo Džeroski, Ljupčo Todorovski, "General meta-model framework for surrogate-based numerical optimization", V: *Discovery science: 20th International Conference, DS 2017, Kyoto, Japan, October 15-17, 2017: proceedings*, (Lecture notes in artificial intelligence, **10558**), Akihiro Yamamoto, ur., 2017, 51-66. [COBISS.SI-ID 30845479]
 22. Matej Martinc, Iza Škrjanec, Katja Zupan, Senja Pollak, "PAN 2017: author profiling - gender and language variety prediction: notebook for PAN at CLEF 2017", V: *CLEF 2017: CLEF 2017 Working notes, Conference and Labs of the Evaluation Forum, Dublin, Ireland, September 11-14, 2017*, (CEUR workshop proceedings, **1866**), Linda Cappellato, ur., 2017, 78. [COBISS.SI-ID 30760743]
 23. Vanja Mileski, Sašo Džeroski, Dragi Kocev, "Predictive clustering trees for hierarchical multi-target regression", V: *Advances in intelligent data analysis XVI: 16th International Symposium on Intelligent Data Analysis, IDA 2017, London, UK, October 26-28, 2017: proceedings*, (Lecture notes in computer science, **10584**), Niall M. Adams, ur., Allan Tucker, ur., David Weston, ur., 2017, 223-234. [COBISS.SI-ID 30883879]
 24. Biljana Mileva-Boshkoska, Dragana Miljković, Anita Valmarska, Dimitros Gatsios, George Rigas, Spyros Konitsiotis, Kostas M. Tsiouris, Dimitrios I. Fotiadis, Marko Bohanec, "A state-transition decision support model for medication Change of Parkinson's disease patients", V: *Slovenska konferenca o umetni inteligenci: zbornik 20. mednarodne multikonference Informacijska družba - IS 2017*, 9.-13. oktober 2017, [Ljubljana, Slovenia]: zvezek A: *proceedings of the 20th International Multiconference Information Society - IS 2017*, 9th-13th October, 2017, Ljubljana, Slovenia: volume A, Mitja Luštrek, ur., Rok Piltaver, ur., Matjaž Gams, ur., Ljubljana, Institut Jožef Stefan, 2017, 63-66. [COBISS.SI-ID 30857255]
 25. Stevanche Nikoloski, Dragi Kocev, Sašo Džeroski, "Decomposition of the output space in multi-label classification using feature ranking", V: *New frontiers in mining complex patterns: Sixth edition of the International Workshop NFMCP 2017 in conjunction with ECML-PKDD 2017, September 18-22, 2017, Skopje, Macedonia*, [S. l., s. n.], 2017. [COBISS.SI-ID 30861863]
 26. Giorgi Peev, Nikola Simidjievski, Sašo Džeroski, "Modeling of dynamical systems: a survey of tools and a case study", V: *Slovenska konferenca o umetni inteligenci: zbornik 20. mednarodne multikonference Informacijska družba - IS 2017*, 9.-13. oktober 2017, [Ljubljana, Slovenia]: zvezek A: *proceedings of the 20th International Multiconference Information Society - IS 2017*, 9th-13th October, 2017, Ljubljana, Slovenia: volume A, Mitja Luštrek, ur., Rok Piltaver, ur., Matjaž Gams, ur., Ljubljana, Institut Jožef Stefan, 2017, 15-18. [COBISS.SI-ID 30853671]
 27. Matej Petković, Sašo Džeroski, Dragi Kocev, "Feature ranking for multi-target regression with tree ensemble methods", V: *Discovery science: 20th International Conference, DS 2017, Kyoto, Japan, October 15-17, 2017: proceedings*, (Lecture notes in artificial intelligence, **10558**), Akihiro Yamamoto, ur., 2017, 171-185. [COBISS.SI-ID 30862887]
 28. Matej Petković, Jovan Tanevski, Aleš Maver, Lovro Vidmar, Borut Peterlin, Sašo Džeroski, "Comparison of feature ranking approaches for discovery of rare genetic variants related to multiple sclerosis", V: *Slovenska konferenca o umetni inteligenci: zbornik 20. mednarodne multikonference Informacijska družba - IS 2017*, 9.-13. oktober 2017, [Ljubljana, Slovenia]: zvezek A: *proceedings of the 20th International Multiconference Information Society - IS 2017*, 9th-13th October, 2017, Ljubljana, Slovenia: volume A, Mitja Luštrek, ur., Rok Piltaver, ur., Matjaž Gams, ur., Ljubljana, Institut Jožef Stefan, 2017, 11-14. [COBISS.SI-ID 30853415]
 29. Senja Pollak, Ana Marija Sobočan, "Corpus-assisted discourse analysis of family legislation changes in Slovene media", V: *9th International Corpus Linguistics Conference, 25 - 28 July 2017, Birmingham, GB: extended abstracts*, Birmingham, University of Birmingham, 2017, 765-769. [COBISS.SI-ID 30666023]
 30. Damjan Popič, Darja Fišer, "Fear and loathing on Twitter: attitudes towards language", V: *Proceedings of the 5th Conference on CMC and Social Media Corpora for the Humanities (cmccorpora17)*, Egon W. Stemle, ur., Ciara R. Wigham, ur., Bolzano, [S. n.], 2017, 61-64. [COBISS.SI-ID 65389666]

31. Andraž Repar, Senja Pollak, "Good examples for terminology databases in translation industry", V: *Electronic lexicography in the 21st century: proceedings of eLex 2017 Conference, 19-21 September 2017, Leiden, The Netherlands*, Izток Kosem, ur., et al, Brno, Lexical Computing, 2017, 650-661. [COBISS.SI-ID 30771751]
32. Andraž Repar, Senja Pollak, "Ontology-based translation memory maintenance", V: *Slovenska konferenca o umetni inteligenci: zbornik 20. mednarodne multikonference Informacijska družba - IS 2017, 9. - 13. oktober 2017, Ljubljana, Slovenija: zvezek C: proceedings of the 20th International Multiconference Information Society - IS 2017, 9-13 October, 2017, Ljubljana, Slovenia: volume C*, Dunja Mladenec, ur., Marko Grobelnik, ur., Ljubljana, Institut Jožef Stefan, 2017, 19-22. [COBISS.SI-ID 30866215]
33. Tanja Samardžić, Mirjana Starović, Željko Agić, Nikola Ljubešić, "Universal dependencies for Serbian in comparison with Croatian and other slavic languages", V: *Proceedings of the EACL workshop, The 6th Workshop on Balto-Slavic Natural Language Processing, April 4, 2017 Valencia, Spain, Stroudsburg, The Association for Computational Linguistics = ACL, 2017, 39-44*. [COBISS.SI-ID 30897703]
34. Tomaž Stepišnik Perdih, Aljaž Osojnik, Sašo Džeroski, Dragi Koccev, "Option predictive clustering trees for hierarchical multi-label classification", V: *Discovery science: 20th International Conference, DS 2017, Kyoto, Japan, October 15-17, 2017: proceedings*, (Lecture notes in artificial intelligence, **10558**), Akihiro Yamamoto, ur., 2017, 116-123. [COBISS.SI-ID 30862631]
35. Jovan Tanevski, Nikola Simidjievski, Ljupčo Todorovski, Sašo Džeroski, "Process-based modeling and design of dynamical systems", V: *Machine learning and knowledge discovery in databases: European Conference, ECML PKDD 2017 Skopje, Macedonia, September 18 - 22, 2017: proceedings. Part I*, (Lecture notes in artificial intelligence, **10534**), Michelangelo Ceci, ur., Ljupčo Todorovski, ur., 2017, 378-382. [COBISS.SI-ID 31053351]
36. Jovan Tanevski, Nikola Simidjievski, Ljupčo Todorovski, Sašo Džeroski, "Process-based modeling and design of dynamical systems", V: *Machine learning and knowledge discovery in databases: European Conference, ECML PKDD 2017 Skopje, Macedonia, September 18-22, 2017. Part III*, (Lecture notes in artificial intelligence, **10536**), Yasemin Altun, ur., Sašo Džeroski, ur., 2017, 378-382. [COBISS.SI-ID 31157543]
37. Anita Valmarska, Dragana Miljković, Spyros Konitsiotis, Dimitros Gatsios, Nada Lavrač, Marko Robnik Šikonja, "Combining multitask learning and short time series analysis in Parkinson's disease patients stratification", V: *Artificial intelligence in medicine: 16th Conference on Artificial Intelligence in Medicine, AIME 2017, Vienna, Austria, June 21-24, 2017: proceedings*, (Lecture notes in artificial intelligence, **10259**), Annette ten Teije, ur., 2017, 116-125. [COBISS.SI-ID 30673447]
38. Anita Valmarska, Dragana Miljković, Marko Robnik Šikonja, Nada Lavrač, "Multi-view approach to Parkinson's disease quality of life data analysis", V: *New frontiers in mining complex patterns: 5th International Workshop NFMCP 2016, held in conjunction with ECML-PKDD 2016, Riva del Garda, Italy, September 19, 2016: revised selected papers*, Annalisa Appice, ur., Cham, Springer, 2017, 163-178. [COBISS.SI-ID 30673191]
39. Ben Verhoeven, Iza Škrjanec, Senja Pollak, "Gender profiling for slovene twitter communication: the influence of gender marking, content and style", V: *Proceedings of the EACL workshop, The 6th Workshop on Balto-Slavic Natural Language Processing, April 4, 2017 Valencia, Spain, Stroudsburg, The Association for Computational Linguistics = ACL, 2017, 119-125*. [COBISS.SI-ID 30412583]
40. Marcos Zampieri, Shervin Malmasi, Nikola Ljubešić, Preslav Nakov, Ahmed Ali, Jörg Tiedemann, Yves Scherrer, Noëmi Aeppli, "Findings of the VarDial evaluation campaign 2017", V: *Proceedings of the Workshop, Fourth Workshop on NLP for Similar Languages, Varieties and Dialects, (VarDial2017), April 3, 2017, Valencia, Spain, Stroudsburg, Association for Computational Linguistics, 2017, 1-15*. [COBISS.SI-ID 30897191]
3. Tomaž Erjavec, Darja Fišer, Nikola Ljubešić, Nataša Logar, Vesna Mikolič, "The expansion of the Gigafida corpus: internet content", V: *Dictionary of modern Slovene: problems and solutions*, (Book series Prevodoslovje in uporabno jezikoslovje), Vojko Gorjanc, ur., et al, 1st ed., e-ed., Ljubljana, Ljubljana University Press, Faculty of Arts, 2017, 120-138. [COBISS.SI-ID 63994722]
4. Tomaž Erjavec, Peter Holozan, Nikola Ljubešić, "Language technologies and corpus encoding", V: *Dictionary of modern Slovene: problems and solutions*, (Book series Prevodoslovje in uporabno jezikoslovje), Vojko Gorjanc, ur., et al, 1st ed., e-ed., Ljubljana, Ljubljana University Press, Faculty of Arts, 2017, 140-154. [COBISS.SI-ID 42764333]
5. Mark A. Finlayson, Tomaž Erjavec, "Overview of annotation creation: processes and tools", V: *Handbook of linguistic annotation*, Nancy M. Ide, ur., James Pustejovsky, ur., James Pustejovsky, Amsterdam, Springer, 2017, 167-192. [COBISS.SI-ID 30613799]
6. Darja Fišer, Jaka Čibej, "Crowdsourcing workflows in lexicography", V: *Dictionary of modern Slovene: problems and solutions*, (Book series Prevodoslovje in uporabno jezikoslovje), Vojko Gorjanc, ur., et al, 1st ed., e-ed., Ljubljana, Ljubljana University Press, Faculty of Arts, 2017, 230-244. [COBISS.SI-ID 63997794]
7. Darja Fišer, Jaka Čibej, "The potential of crowdsourcing in modern lexicography", V: *Dictionary of modern Slovene: problems and solutions*, (Book series Prevodoslovje in uporabno jezikoslovje), Vojko Gorjanc, ur., et al, 1st ed., e-ed., Ljubljana, Ljubljana University Press, Faculty of Arts, 2017, 212-228. [COBISS.SI-ID 63996770]
8. Darja Fišer, Tomaž Erjavec, Nikola Ljubešić, "The compilation, processing and analysis of the JANES corpus of Slovene user-generated content", V: *Corpus de communication médiée par les réseaux: construction, structuration, analyse*, (Collection Humanités numériques), Ciara R. Wigham, ur., Gudrun Ledegen, ur., Paris, L'Harmattan, cop. 2017, 125-138. [COBISS.SI-ID 64650338]
9. Maja Miličević, Nikola Ljubešić, Darja Fišer, "Birds of a feather don't quite tweet together: an analysis of spelling variation in Slovene, Croatian and Serbian twitterese", V: *Investigating computer-mediated communication: corpus-based approaches to language in the digital world*, (Book series Translation studies and applied linguistics), Darja Fišer, ur., Michael Reißwenger, ur., 1st ed., Ljubljana, University Press, Faculty of Arts, 2017, 14-43. [COBISS.SI-ID 65266274]
10. Friedrich Recknagel, Dragi Koccev, Cao Hongqin, Christina Castelo Branco, Ricardo Minoti, Sašo Džeroski, "Inferential modelling of population dynamics", V: *Ecological informatics: data management and knowledge discovery*, Friedrich Recknagel, ur., William K. Michener, ur., 3rd ed., Cham, Springer, 2017, 161-188. [COBISS.SI-ID 30863911]
11. Andrejka Žejn, Tomaž Erjavec, Matija Ogrin, "Poljanski rokopis: Baročno besedilno izročilo o Jezusovem življenju v elektronski izdaji", V: *Rokopisi slovenskega slovstva od srednjega veka do moderne*, (Obdobja, 36), Aleksander Bjelčević, ur., Matija Ogrin, ur., Urška Perenič, ur., 1. natis, Ljubljana, Znanstvena založba Filozofske fakultete, 2017, 111-118. [COBISS.SI-ID 65662562]

ZNANSTVENA MONOGRAFIJA

1. Pavle Boškosi, Andrej Debenjak, Biljana Mileva-Boshkoska, *Fast electrochemical impedance spectroscopy: as a statistical condition monitoring tool*, (Springer briefs in applied sciences and technology), Cham, Springer, 2017. [COBISS.SI-ID 30534439]
2. Blaž Podpečan, Vid Podpečan, *Govorica tišine: nagrobnik kot vir podatkov*, (Archaeologia historica Slovenica, 7), 1. izd., Ljubljana, Znanstvena založba Filozofske fakultete, 2017. [COBISS.SI-ID 290659840]

MENTORSTVO

1. Jože Bučar, *Klasifikacija spletnih besedil na osnovi izraženosti sentimenta*: doktorska disertacija, Novo Mesto, 2017 (mentor Janez Povh; somentor Martin Žnidaršič). [COBISS.SI-ID 2048474131]
2. Jan Kralj, *Analiza heterogenih informacijskih omrežij za semantično rudarjenje podatkov*: doktorska disertacija, Ljubljana, 2017 (mentor Nada Lavrač; somentor Marko Robnik Šikonja). [COBISS.SI-ID 292862464]
3. Janez Kranjc, *Spletni delotoki za rudarjenje podatkov v oblaku*: doktorska disertacija, Ljubljana, 2017 (mentor Nada Lavrač; somentor Marko Robnik Šikonja). [COBISS.SI-ID 290659584]
4. Jurica Levatič, *Polnadzorovano učenje za napovedovanje strukturiranih vrednosti*: doktorska disertacija, Ljubljana, 2017 (mentor Sašo Džeroski; somentor Dragi Koccev). [COBISS.SI-ID 290735616]

SAMOSTOJNI ZNANSTVENI SESTAVEK ALI POGLAVJE V MONOGRAFSKI PUBLIKACIJI

1. Kaja Dobrovoljc, Simon Krek, Tomaž Erjavec, "The Sloleks morphological lexicon and its future development", V: *Dictionary of modern Slovene: problems and solutions*, (Book series Prevodoslovje in uporabno jezikoslovje), Vojko Gorjanc, ur., et al, 1st ed., e-ed., Ljubljana, Ljubljana University Press, Faculty of Arts, 2017, 42-63. [COBISS.SI-ID 63993954]
2. Tomaž Erjavec, "MULTEXT-East", V: *Handbook of linguistic annotation*, Nancy M. Ide, ur., James Pustejovsky, ur., James Pustejovsky, Amsterdam, Springer, 2017, 441-462. [COBISS.SI-ID 30614055]