DEPARTMENT OF KNOWLEDGE TECHNOLOGIES E-8

Department of Knowledge Technologies performs research in advanced information technologies aimed at acquiring, storing and managing knowledge to be used in the development of an information- and knowledge-based society. Established areas of our work include intelligent data analysis (machine learning, data mining, and knowledge discovery in databases), language technologies and computational linguistics, computational creativity, decision support and knowledge management. In addition to research in knowledge technologies, we are also developing applications in environmental sciences and management, agronomy, medicine, biomedicine and bioinformatics, economics, finance and marketing. The department is also a recognised centre for linguistic research and digital humanities.

Head:
Prof. Nada Lavrač

In 2016 we were involved in twelve national projects, eight EU FP7 and four Horizon 2020 projects, one COST action, two bilateral projects, one infrastructure project and four industry projects. The department hosted nine junior researchers working towards their PhDs.

In the area of **intelligent data analysis and data mining** we have developed several new methods and used them in a number of application domains. In the area of semantic data mining we have developed a method for explaining mixture models through semantic pattern mining and banded matrix visualization, published in the respected Machine Learning journal and a method for network-ranking-assisted semantic data mining. In the area of heterogeneous network analysis we developed new search-heuristics-based text mining. We continued work in the area of multi-view clustering using multi-target predictive clustering trees, with a similar method applied to clustering of Alzheimer's disease patients. A new web platform for text mining and natural language processing TextFlows was developed, together with complex language processing workflows, and presented in the Science of Computer Programming journal. In text mining we continued research in bisociative cross-domain link discovery. Within the PD_manager project, which we coordinate at JSI, we developed a multi-view methodology for determining groups of patients with similar symptoms and detecting patterns of medication changes that lead to the improvement or decline of patients' quality of life.

Our team won the ESA's Mars Express Power Challenge, designing the most accurate solution for predicting a space probe's power consumption.

We have developed new methods for the automated modelling of dynamic systems, using both data and domain knowledge, and have applied them to problems in the areas of environmental and life sciences. These include methods for learning ensembles of deterministic process-based models and methods for learning stochastic process-based models. We have adapted the methods for learning process-based models for the task of designing dynamic biological systems and demonstrated their utility on several problems from synthetic biology.

We are coordinating the FP7 FET project MAESTRA (Learning from Massive, Incompletely Annotated, and Structured Data), which focuses on analysing data that may be complex in several ways. Within MAESTRA we have developed many new methods for structured output prediction in a batch or streaming setting. These include methods for multi-target regression, such as learning option trees in batch setting, as well as learning trees and tree ensembles in streaming setting. The latter have been used to perform multi-label classification on data streams. We have also developed methods for feature ranking for different types of structured outputs. Finally, we have proposed a method to organize the labels in multi-label classification into a hierarchy, which improves the predictive performance.

We have used methods of predictive clustering for the analysis of various aspects of the biology of extremophilic fungi, which can act as opportunistic human pathogens. We have also used these methods for the identification of multi-output dynamic systems in batch and streaming settings. Finally, we used methods for multi-target regression in the ESA Mars Express Power Challenge, where the task was to predict the power consumption of 33 different lines of the thermal subsystem for the Mars Express orbiter.

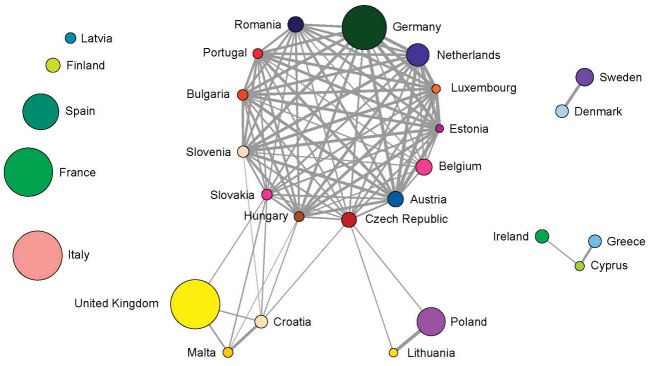

Figure 1: A network of the EU countries, linked by tweet sharing. The size of a node is proportional to the number of EP members from that country. The network contains one large connected component consisting of 18 countries, with a core of 13 countries around Germany that are densely connected.

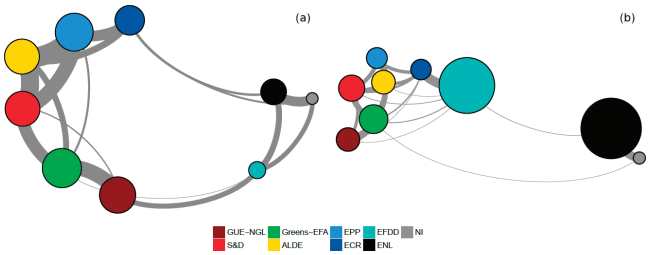

Figure 2: Networks of roll-call votes and retweets. (A) Co-voting agreement within and between political groups in the EP. (B) Average retweets within and between political groups.

In the context of the FET Flagship HBP (Human Brain Project) project, we are developing new data-mining methods and applying them to discover biological signatures of neurodegenerative diseases, such as Alzheimer's and Parkinson's. A new multi-view clustering method was developed and used for the identification of gender-specific biomarkers for Alzheimer's disease. Several methods for redescription mining, based on predictive clustering, were developed and applied to data on Alzheimer's disease. Finally, we have applied predictive clustering trees to relate biomarkers and clinical scores for Parkinson's disease.

In collaboration with the French Institute ARVALIS (Institut du végétal), we launched the project BIODIV (Understanding and managing biodiversity in agricultural ecosystems by data mining and decision support). We organized a workshop where we explored the current state of data collection and the modelling of habitats and population dynamics for auxiliary species that perform biological control of pests (such as aphids) in cereal crops. We performed extensive and complex data pre-processing, grouping organisms into functional categories and linking sampling sites to their environmental characteristics, thus getting ready to apply data mining and decision modelling.

In the area of **text and web mining, and heterogeneous information network analysis**, we successfully completed (with an excellent evaluation) the MULTIPLEX project (Foundational Research on MULTIlevel comPLEX networks and systems). We continued work on an existing FP7 project SIMPOL (Financial Systems Simulation and Policy Modelling) and a H2020 FET project DOLFINS (Distributed Global Financial Systems for Society). We also started a new national project FORMICA (Influence of formal and informal corporate communications on capital markets).

The NAO robot talking through the WHIMBOT interface, developed within the WHIM project, attracted a lot of attention at this year's International fair of crafts and entrepreneurship in Celje.

The main emphasis of the research in this area is combining text mining, network analysis and sentiment analysis to reveal and highlight underlying properties in different domains. The main sources of data that we analyse are social media (Twitter, Facebook), online news, and other relevant data (e.g., voting in the European Parliament, stock prices, results of elections and referendums).

We collected tweets from the members of the European Parliament (EP), constructed retweet networks, and identified political and national communities. We also analysed the voting of the current, eighth EP, and compared the relation between the co-voting and retweeting. The results show that retweeting is very similar to co-voting, but also that the right-wing political groups are much more active and coherent in social media.

Figure 3: Winning team of the ESA's Mars Express Power Challenge visiting ESA Mission Control Centre (ESOC) in Darmstadt.

We developed methodology and tools to analyse Twitter sentiment (or stance) in different languages and for different topics. Our approach relies on a large set of manually annotated tweets by domain experts, and results in high-quality domain- and language-specific classification models. We showed how to estimate an upper bound of performance for a classification model and how to monitor the annotation process.

Finally, we applied the above-developed methodology to a highly relevant European use-case: the UK referendum on Brexit. We developed a Brexit-specific Twitter stance classification model and used it to monitor the public mood before and during the referendum. We also identified the most influential Twitter users and noted that the proponents of Brexit were considerably more active than the remain camp.

In the area of **decision support** our long-term goal is to develop methods and techniques of decision modelling, support them with software and integrate them with data-mining systems. In 2016 we published the results of applying these methods in two areas: assessment of long-term sustainability of electric energy production in Slovenia and assessment of potential presence of genetically modified crops in food and feed based on traceability data (Project DECATHLON). In each of these areas we developed a corresponding decision-support system. We also successfully used our qualitative multi-criteria modelling method DEX for the assessment of anthropogenic activities on the ecology of the Fifth Triglav Lake in the Julian Alps and for the evaluation of business process simulation software. We investigated and compared three methods for approximating incompletely defined utility functions with the DEX method. In the framework of the European project HeartMan we were developing decision models and systems for supporting patients having Congestive Heart Failure. For the European project PD_manager we have developed several decision-support models based on the expert knowledge of the clinical project partners, which reflect current knowledge in the Parkinson's disease domain regarding the medication changes, but had not been previously formalized. Regarding software

tools, in addition to updating the existing DEXi model-development program and JDEXi library, we developed the DEXi HTML Evaluator, a new software for the evaluation of decision alternatives using DEX models, which runs on web pages that require almost no programming to set up.

In the area of **language technologies and digital humanities** we work on producing language resources and methods to annotate text with linguistic information, with a focus on the Slovene language. In 2016 we helped to produce the digital text-critical edition of the 18th century Kapla passion play and worked on methods to modernise historical Slovene language, as well as standardising Slovene as found on the web. We have converted the ssj500k treebank into the Universal Dependencies framework and made the resulting treebank available in the scope of the UD project.

Vid Podpečan received the Prometheus of Science award for excellence in communication for the year 2016 from the Slovenian Science Foundation.

We continued work on the national research project JANES "Resources, Tools and Methods for Research of Non-standard Internet Slovene", where we compiled the penultimate version of the JANES corpus of user-generated internet Slovene, which contains over 160 million words of tweets, news comments, forums, blogs and Wikipedia talk pages and researched ways in which to make such corpora openly available. We developed a method for restoring diacritics which are often missing in internet texts, developed a new part-of-speech tagger and lemmatiser, made our tools available as web applications, performed and analysed automatic sentiment labelling on internet texts, developed a method to detect semantic shifts in Slovene tweets, to distinguish private from corporate users, and performed analytics over them. We also studied the global use of emojis in tweets as well as sociolinguistic aspects, regarding the relation between the topic and gender of the user as well as the gender-representation on the web. We worked on the manual annotation of user-generated internet texts, concentrating on normalisation, part-of-speech tagging, lemmatisation and syntactic annotation with the motivation to produce training sets for machine-learning annotation tools for (non-standard) Slovene. We started work on a dictionary of Twitter-specific lexis and undertook several linguistic analyses on various aspects of user-generated internet texts, such as the use of the comma.

Figure 4: Our NAO WHIMBOT attracted a lot of attention at the International fair of crafts and entrepreneurship in Celje.

We started work on our new national research project KAS "Slovene scientific texts: resources and description", where we produced the prototype version of the KAS corpus, which contains almost one billion words of scientific texts gathered from the digital libraries of Slovene universities and, on the basis of this corpus, performed an analysis on the use of the passive in Slovene scientific writing. As preparation on terminology extraction, we constructed a term grammar for the Sketch Engine corpus analysis platform.

In the scope of the industrial project TermIolar with the Slovene translation company Iolar, we developed a system for terminology management, integrating automatic terminology extraction from translation memories and existing terminological databases.

We lead the Slovene research infrastructure CLARIN.SI, which provides easy publication and sustainable access to digital language data for scholars in the humanities and social sciences. In 2016 the CLARIN.SI repository for language resources and tools was recognised as a CLARIN Center B, certifying it as a stable and trusted repository that conforms to the CLARIN technical and organisational standards. In this year the number of deposited resources in the repository more than doubled, reaching 47 at the end of the year, with the most interesting additions being a transcribed speech corpus of Slovene public lectures and a corpus of parliamentary debates from the years 1990–1992. In 2016 CLARIN.SI has also organised the international workshop "Multilingual corpus annotation tools: development and integration" and supported several conferences and summer schools.

We collaborated in the work of the Slovene Institute for Standardization as the Slovene representatives in ISO/TC37/SC4 (Terminology and Other Language and Content Resources / Language Resources Management) by reviewing, translating and approving Slovene standards from this field. We also continue to serve as technical editors for the on-line Slovene Biographic Lexicon.

In 2016 we organised the 10th biennial conference on Language Technologies, which was this year extended with the theme of Digital Humanities. The conference, which took place at the Arts Faculty of the University of Ljubljana, lasted three days with parallel sessions, and comprised five invited lectures, a round table on terminology, a student session, and 47 regular contributions.

In the field of **computational creativity** we continued to develop new prototype solutions using our tool ConCreTeFlows, which is based on our platform ClowdFlows. Within the FP7 project CONCRETE (Concept Creation Technologies) we prepared a workflow for the multimodal mixing of concepts, which also served for the demonstra-

tion of ConCreTeFlows and its ability to support teamwork and the use of a variety of program elements in a single workflow. Presentation of the system with this demonstration at the International Conference on Computational Creativity ICCC 2016 was met with a positive response from the research community. We analysed the role of optimality principles in different conceptual blending systems and proposed the conceptualization of computational creativity. Within FP7 project WHIM (The What-If Machine) we conducted the final experiments of human appreciation modelling of computer generated what-if ideas and presented the system for scientific question generation. This year, the project WHIM achieved a remarkable achievement for the computational creativity research community: on the basis of computer-generated ideas from the WHIM system and other computational creativity approaches, a musical was made and presented on stage in London. In relation to that, a documentary was filmed about this venture and the technologies that were used to achieve it. In 2016, the projects of FP7: Prosecco, WHIM (The What-If Machine) and ConCreTe (Concept Creation Technologies) were finished and the results of the projects, as well as the wider area of computational creativity, were presented at invited talk of the TTT conference TTT 2016.

Sašo Džeroski was elected to the European Academy of Sciences (Academia Europaea).

Some outstanding publications in the last year

1. Adhikari, P. R., Vavpetič, A., Kralj, J., Lavrač, N., Hollmén, J. Explaining mixture models through semantic pattern mining and banded matrix visualization. *Machine learning*, ISSN 0885-6125. 2016, 37 pp., doi: 10.1007/s10994-016-5550-3.
2. Bohanec, M., Mileva-Boshkoska, B., Prins, T. W., Kok, E. SIGMO: a decision support System for Identification of genetically modified food or feed products. Food control, ISSN 0956-7135. 2016, vol. 71, pp. 168-177, doi: 10.1016/j.foodcont.2016.06.032.
3. Bohanec, M., Trdin, N., Kontić, B. 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, 2016, 15 pp., doi: 10.1007/s10100-016-0457-4.
4. Cherepnalkoski, D., Karpf, A., Mozetič, I., Grčar, M. Cohesion and coalition formation in the European Parliament: Roll-call votes and Twitter activities, PLoS ONE 11(11): e0166586, doi: 10.1371/journal.pone.0166586, 2016.
5. Mozetič, M. Grčar, J. Smailović. Multilingual Twitter sentiment classification: The role of human annotators, PLoS ONE 11(5): e0155036, doi: 10.1371/journal.pone.0155036, 2016.
6. Osojnik, A., Panov, P., Džeroski, S. Multi-label classification via multi-target regression on data streams. *Machine learning*, ISSN 0885-6125. In press, 2016, 26 pp., doi: 10.1007/s10994-016-5613-5.
7. Perovšek, M., Kranjc, J., Erjavec, T., Cestnik, B., Lavrač, N. TextFlows : a visual programming platform for text mining and natural language processing. *Science of computer programming*, ISSN 0167-6423, 2016, vol. 121, str. 128-152, doi: 10.1016/j.scico.2016.01.001.
8. Scherrer, Y., Erjavec, T. Modernising historical Slovene words. *Natural language engineering*, ISSN 1351-3249, 2016, vol. 22, no. 6, pp. 881-905, doi: 10.1017/S1351324915000236.
9. Tanevski, J., Todorovski, L., Džeroski, S. Process-based design of dynamical biological systems. Scientific reports, ISSN 2045-2322, 2016, vol. 6, pp. 34107-1-34107-13, doi: 10.1038/srep34107.
10. Tanevski, J., Todorovski, L., Džeroski, S. Learning stochastic process-based models of dynamical systems from knowledge and data. *BMC systems biology*, ISSN 1752-0509, 2016, vol. 10, pp. 30-1-30-17, doi: 10.1186/s12918-016-0273-4.

Awards and appointments

1. Matej Petković was awarded the Prešeren Prize of the Faculty of Mathematics and Physics of the University of Ljubljana for his master's thesis titled "Evaluation of the importance of continuous features with the ReliefF method", written under the supervision of Sašo Džeroski.
2. Sašo Džeroski, corresponding member of the Macedonian Academy of Arts and Sciences, was elected for a full member of the European Academy of Sciences (Academia Europaea).
3. Vid Podpečan received the Prometheus of Science award for excellence in communication for the year 2016 from the Slovenian Science Foundation for the all-round promotion of robotics, especially humanoid robot use among young people.
4. Special recognition for successful cooperation with the Committee on Science and Technology of the Chamber of craft and small business of Slovenia and for an interesting presentation of the Department of Knowledge Technologies of the Jožef Stefan Institute at the International fair of crafts and entrepreneurship in Celje.

5. Martin Breskvar, Dragi Kocev, Jurica Levatić, Aljaž Osojnik, Matej Petković, Nikola Simidjievski and Bernard Ženko won the ESA's Mars Express Power Challenge, designing the most accurate solution for predicting the space probe's power consumption.

Organization of conferences, congresses and meetings

1. Workshop CLARIN "Multilingual corpus annotation tools: development and integration", Ljubljana, 10.–11. 11. 2016
2. "Summer school on mining big and complex data", Ohrid, Macedonia, 4.–8. 9. 2016
3. Final meeting of EU project "Concept Creation Technology" - ConCreTe, Bled, 14.–16. 9. 2016.

INTERNATIONAL PROJECTS

1. BIODIV - Understanding and Managing Biodiversity in Agricultural Ecosystems by Data Mining and Decision Support
   Prof. Sašo Džeroski
   Arvalis - Institut Du Végétal
2. 7FP - PROSECCO; Promoting the Scientific Exploration of Computational Creativity
   Prof. Nada Lavrač
   European Commission
3. 7FP - ConCreTe; Concept Creation Technology
   Prof. Nada Lavrač
   European Commission
4. 7FP - WHIM; The What-If Machine
   Prof. Nada Lavrač
   European Commission
5. 7FP - DECATHLON; Development of Cost efficient Advanced DNA-based methods for specific Traceability issues and High Level On-site applicatioNs
   Prof. Marko Bohanec
   European Commission
6. 7FP - SIMPOL; Financial Systems Simulation and Policy Modelling
   Prof. Igor Mozetič
   European Commission
7. 7FP - MULTIPLEX; Foundational Research on Multilevel Complex Networks and Systems
   Prof. Igor Mozetič
   European Commission
8. 7FP - HBP; The Human Brain Project
   Prof. Sašo Džeroski
   European Commission
9. 7FP - MAESTRA; Learning from Massive, Incompletely Annotated, and Structured Data
   Prof. Sašo Džeroski
   European Commission
10. PARSEME: PARSing and Multi-Word Expressions. Towards Linguistic Precision and Computational Efficiency in Natural Language Processing
    Prof. Tomaž Erjavec
    Cost Office
11. H2020 - DOLFINS; Distributed Global FInancial Systems for Society
    Prof. Igor Mozetič
    European Commission
12. H2020 - HBP SGA1; Human Brain Project Specific Grant Agreement 1
    Prof. Sašo Džeroski
    European Commission
13. H2020 - PD_manager; Mhealth Platform for Parkinson's Disease Management
    Dr. Dragana Miljković
    European Commission
14. H2020 - HeartMan; Personal Decison Support System for Heart Failure Management
    Prof. Marko Bohanec
    European Commission
15. Mining Complex Data in Environmental and Life Sciences
    Prof. Sašo Džeroski
    Slovenian Research Agency
16. Mining Heterogeneous Linked Biomedical Data
    Prof. Nada Lavrač
    Slovenian Research Agency
17. Semantic Role Labeling in Slovene and Croatian
    Prof. Tomaž Erjavec
    Slovenian Research Agency

RESEARCH PROGRAM

1. Knowledge Technologies
   Prof. Nada Lavrač

R & D GRANTS AND CONTRACTS

1. Slovenian Literature in Unknown Early Modern Manuscripts: Information-Technology Aided Analyses and Scholarly Editions
   Prof. Tomaž Erjavec
2. Integrative research of sexual dimorphism evolution
   Prof. Sašo Džeroski
3. Resources, Tools and Methods for the Research of Nonstandard Internet Slovene
   Prof. Tomaž Erjavec
4. Slovene scientific texts: resources and description
   Prof. Tomaž Erjavec
5. Analysis of heterogeneous information networks for knowledge discovery in life-sciences
   Prof. Nada Lavrač
6. Molecular bases of interactions among the grapevine and phytoplasmal causing agents of the grapevine yellows diseases
   Prof. Nada Lavrač
7. Development of a multimethod approach to study wildlife behavior: investigating humanbear conflicts in contrasting landscapes of Europe
   Prof. Sašo Džeroski
8. Influence of formal and informal corporate communications on capital markets
   Dr. Senja Pollak
9. Development and applications of new semantic data mining methods in life sciences
   Prof. Nada Lavrač
10. Machine Learning for Systems Sciences
    Prof. Sašo Džeroski
11. Forbidden Books in the Slovenian Lands in the Early Modern Period
    Prof. Tomaž Erjavec
12. Structured output prediction with applications in sustainable agricultural production
    Prof. Sašo Džeroski
13. ReLDI - Regional Linguistic Data Initiative
    Dr. Nikola Ljubešić
    University Of Zurich, Urpp Language And Space
14. CLARIN Workshop: „Multilingual Corpus Annotation Development and Integration", IJS, Ljubljana, Slovenia; 10.11.-11.11.2016
    Prof. Tomaž Erjavec
    Clarin Eric
15. Methods of digital ethnography and monitoring of experiences and the development of global competencies of Erasmus students
    Dr. Martin Žnidaršič
    CMEPIUS

NEW CONTRACTS

1. Development of a Prototype program solution for support of semi-automatic extraction and management of monolingual and multilingual corpora
   Dr. Senja Pollak
   Iolar d. o. o.

VISITORS FROM ABROAD

1. Dario Stojanovski, Ss. Cyrill and Methodius, Faculty of Computer Science and Engineering, Skopje, Macedonia, 18.1.-7.2. 2016
2. Dr. Gjorgi Strezoski, Ss. Cyrill and Methodius, Faculty of Computer Science and Engineering, Skopje, Macedonia, 18.1.-7.2. 2016
3. Dr. Gjorgi Madjarov, Ss. Cyrill and Methodius, Faculty of Computer Science and Engineering, Skopje, Macedonia, 31.1.-7.2.2016
4. Dr. Ivica Dimitrovski, Ss. Cyrill and Methodius, Faculty of Computer Science and Engineering, Skopje, Macedonia, 31.1.-7.2.2016

5. Andreas Karpf, UPIP - Centre d'Economie de la Sorbonne, Université Paris, Paris, France, 20.2.-31.3.2016
6. Matej Mihelčić, Institut Ruđer Bošković, Zagreb, Croatia, 3.3.2016
7. Matija Piškorec, Institut Ruđer Bošković, Zagreb, Croatia, 3.3.2016
8. dr. Larisa Soldatova, Brunel University, London, Great Britain, 21. 3. 2016
9. prof. dr. Ross King, University of Manchester, Manchester, Great Britain, 21.3.2016
10. Fabiana Zollo, IMT, School for Advanced studies Lucca, Lucca, Italy, 9.4.–1.7.2016.
11. Amilcar Cardoso, Coimbra University, Coimbra, Portugal, 13.–15. 4. 2016
12. Ana Oliveira Alves, Coimbra University, Coimbra, Portugal, 13.–15. 4. 2016
13. João Carlos Gonçalves, Coimbra University, Coimbra, Portugal, 13.–15. 4. 2016
14. dr. Tomislav Šmuc, Institut Ruđer Bošković, Zagreb, Croatia, 3.3.2016 and 9.–13. 5. 2016
15. prof. dr. Filip Železny, Department of Computer Science, Faculty of Electrical Engineering, Czech Technical University in Prague, Prague, Czech Republic, 8.–10. 5. 2016
16. Jožef Mišutka, Karlova univerza v Pragi, Prague, Czech Republic, 11.–16. 4. 2016
17. dr. Barry Hardy, Douglas Connect, Basel, Switzerland , 3.–6. 5. 2016
18. prof. dr. Ljupčo Kocarev, Macedonian Academy of Sciences and Arts, Skopje, Macedonia, 22.–24. 5. 2016
19. Ivan Stojković, Temple University, Philadelphia, USA, 20.–22.7.2016
20. Dr. Daniel Zeman, Karlova Univerza, Prague, Czech Republic, 28.9.–1. 10. 2016
21. dr. Petru Juelu Henrichsenu, Copenhagen Business School, Copenhagen, Denmark, 29. 9.– 1. 10. 2016
22. Filip Petkovski, Zagreb, Croatia, 26.–28.9.2016
23. Vera Veleva, BILSP, Bulgaria, 25. 10. 2016
24. Yordan Dimitrov, BILSP, Bulgaria, 25. 10. 2016
25. Ben Verhoeven, University of Antwerp, Antwerp, Belgium, 31. 10.-26. 11. 2016
26. prof. dr. Bogdan Draganski Centre hospitalier universitaire vaudois, Luzana, Switzerland, 13.–15. 11. 2016
27. dr. Florence Leprince, ARVALIS-Institut du végétal, Pau, France, 24.–25. 11. 2016
28. dr. Markus Schedl, Department of Computational Perception, Johannes Kepler University (JKU) Linz, Austria, 15. 12. 2016

STAFF

Researchers
1. Prof. Marko Bohanec
2. Prof. Bojan Cestnik*
3. Prof. Marko Debeljak
4. Prof. Sašo Džeroski
5. Prof. Tomaž Erjavec
6. Dr. Darja Fišer*
7. Dr. Dragi Kocev
8. **Prof. Nada Lavrač, Head**
9. Prof. Zoran Levnajić*
10. Prof. Igor Mozetič
11. Prof. Ljupčo Todorovski*
12. Prof. Tanja Urbančič*
13. Asst. Prof. Martin Žnidaršič

Postdoctoral associates
14. Dr. Darko Aleksovski
15. *Dr. Darko Cherepnalkoski, left 01.11.16*
16. Dr. Miha Grčar
17. Dr. Petra Kralj Novak
18. Dr. Nikola Ljubešić
19. Asst. Prof. Biljana Mileva Boshkoska
20. Dr. Dragana Miljković
21. Asst. Prof. Panče Panov
22. Dr. Vid Podpečan
23. Dr. Senja Pollak
24. Dr. Nikola Simidjievski
25. *Dr. Borut Sluban, left 01.10.16*
26. Dr. Jasmina Smailović
27. Dr. Jovan Tanevski
28. Dr. Aneta Trajanov
29. Dr. Anže Vavpetič
30. Dr. Vedrana Vidulin
31. Asst. Prof. Bernard Ženko

Postgraduates
32. Martin Breskvar, B. Sc.
33. Jan Kralj, B. Sc.
34. Janez Kranjc, B. Sc.
35. Jurica Levatić
36. Aljaž Osojnik
37. *Dr. Matic Perovšek, left 01.07.16*
38. Matej Petković, B. Sc.
39. Tomaž Stepišnik Perdih, B. Sc.
40. Anita Valmarska, B. Sc.
41. Katja Zupan, B. Sc.

Technical officers
42. Tina Anžič, B. Sc.
43. Milica Bauer, B. Sc.

Technical and administrative staff
44. *Teja Đukić, 01.10.16, transferred to Department K7*

Note:
* part-time JSI member

BIBLIOGRAPHY

Original Article

1. Prem Raj Adhikari, Anže Vavpetič, Jan Kralj, Nada Lavrač, Jaakko Hollmén, "Explaining mixture models through semantic pattern mining and banded matrix visualization", *Mach. learn.*, vol. 105, iss. 1, pp. 3-39, 2016.
2. Darko Aleksovski, Juš Kocijan, Sašo Džeroski, "Ensembles of fuzzy linear model trees for the identification of multi-output systems", *IEEE trans. fuzzy syst.*, vol. 24, no. 4, pp. 916-929, 2016.
3. Marko Bohanec, Biljana Mileva-Boshkoska, Theo W. Prins, Esther Kok, "SIGMO: a decision support System for Identification of genetically modified food or feed products", *Food control*, vol. 71, pp. 168-177, 2016.
4. Marija Brbić, Matija Piškorec, Vedrana Vidulin, Anita Kriško, Tomislav Šmuc, Fran Supek, "The landscape of microbial phenotypic traits and associated genes", *Nucleic acids res.*, vol. 44, no. 21, pp. 10074-10090, 2106.
5. Darko Čerepnalkoski, Andreas Karpf, Igor Mozetič, Miha Grčar, "Cohesion and coalition formation in the European parliament: roll-call votes and twitter activities", *PloS one*, vol. 11, no. 11, pp. 0166586-1-0166586-27, 2016.
6. Darko Čerepnalkoski, Igor Mozetič, "Retweet networks of the European Parliament: evaluation of the community structure", *Appl. netw. sci.*, vol. 1, no. 1, pp. 2-1-2-20, 2016.
7. Nadja Damij, Pavle Boškoski, Marko Bohanec, Biljana Mileva-Boshkoska, "Ranking of business process simulation software tools with DEX/QQ hierarchical decision model", *PloS one*, vol. 11, no. 2, pp. 0148391-1-0148391-16, 2016.
8. Ivica Dimitrovski, Dragi Kocev, Suzana Loskovska, Sašo Džeroski, "Improving bag-of-visual-words image retrieval with predictive clustering trees", *Inf. sci.*, vol. 329, pp. 851-865, 2016.
9. Tomaž Erjavec, Jaka Čibej, Darja Fišer, "Omogočanje dostopa do korpusov slovenskih spletnih besedil v luči pravnih omejitev", In: *Računalniško posredovana komunikacija*, (Slovenščina 2.0, Tematska številka, vol. 4 (2016), no. 2), Darja Fišer, ed., Ljubljana, Trojina, zavod za uporabno slovenistiko, 2016, vol. 4, no. 2, pp. 189-219, 2016.
10. Darja Fišer, Tomaž Erjavec, Nikola Ljubešić, "JANES v0.4: korpus slovenskih spletnih uporabniških vsebin", In: *Računalniško posredovana komunikacija*, (Slovenščina 2.0, Tematska številka, vol. 4 (2016), no. 2), Darja Fišer, ed., Ljubljana, Trojina, zavod za uporabno slovenistiko, 2016, vol. 4, no. 2, pp. 67-99, 2016.
11. Dragan Gamberger, Bernard Ženko, Alexis Mitelpunkt, Netta Shachar, Nada Lavrač, "Clusters of male and female Alzheimer's disease patients in the Alzheimer's Disease Neuroimaging Initiative (ADNI) database", *Brain inform.*, vol. 3, iss. 3, pp. 169-179, 2016.
12. Sašo Jančič, Jens C. Frisvald, Dragi Kocev, Cene Gostinčar, Sašo Džeroski, Nina Gunde-Cimerman, "Production of secondary metabolites in extreme environments: food- and airborne Wallemia spp. produce toxic

metabolites at hypersaline conditions", *PloS one*, vol. 11, no. 12, pp. e0169116-1-e0169116-20, 2016.
13. Sašo Jančič, Polona Zalar, Dragi Kocev, Hans-Josef Schroers, Sašo Džeroski, Nina Gunde-Cimerman, "Halophily reloaded: new insights into the extremophilic life-style of Wallemia with the description of Wallemia hederae sp. nov", *Fungal divers.*, vol. 76, issue 1, pp. 97-118, 2016.
14. Branko Kontić, Marko Bohanec, Davor Kontić, Nejc Trdin, Maruša Matko, "Improving appraisal of sustainability of energy options - a view from Slovenia", *Energy policy*, vol. 90, pp. 154-171, 2016.
15. Zoran Krstanoski, Nadja Kokalj-Vokač, Andreja Zagorac, Boris Pospihalj, Miha Munda, Sašo Džeroski, Rastko Golouh, "TMPRSS2:ERG gene aberrations may provide insight into pT stage in prostate cancer", *BMC Urol*, vol. 16, pp. 1-6, 2016.
16. Marta Macedoni-Lukšič, Tanja Urbančič, Ingrid Petrič, Bojan Cestnik, "Autism research dynamic through ontology-based text mining", *Advances in autism*, vol. 2, no. 3, pp. 131-139, 2016.
17. Gjorgji Madjarov, Dejan Gjorgjevikj, Ivica Dimitrovski, Sašo Džeroski, "The use of data-derived label hierarchies in multi-label classification", *Journal of intelligent information systems*, vol. 47, iss. 1, pp. 57-90, 2016.
18. Maja Miličević, Nikola Ljubešić, "Tviterasi, tviteraši or twitteraši?: producing and analysing a normalised dataset of Croatian and Serbian tweets", In: *Računalniško posredovana komunikacija*, (Slovenščina 2.0, Tematska številka, vol. 4 (2016), no. 2), Darja Fišer, ed., Ljubljana, Trojina, zavod za uporabno slovenistiko, 2016, vol. 4, no. 2, pp. 156-188, 2016.
19. Igor Mozetič, Miha Grčar, Jasmina Smailović, "Multilingual twitter sentiment classification: the role of human annotators", *PloS one*, vol. 11, no. 5, pp. e0155036-1-e0155036-26, 2016.
20. Monika Novak, Polona Zalar, Bernard Ženko, Sašo Džeroski, Nina Gunde-Cimerman, "Yeasts and yeast-like fungi in tap water and groundwater, and their transmission to household appliances", *Fungal ecology*, vol. 20, pp. 30-39, 2016.
21. Aljaž Osojnik, Panče Panov, Sašo Džeroski, "Modeling dynamical systems with data stream mining", *Comput. Sci. Inf. Syst.*, vol. 13, no. 2, pp. 453-473, 2016.
22. Panče Panov, Larisa N. Soldatova, Sašo Džeroski, "Generic ontology of datatypes", *Inf. sci.*, vol. 329, pp. 900-920, 2016.
23. Matic Perovšek, Janez Kranjc, Tomaž Erjavec, Bojan Cestnik, Nada Lavrač, "TextFlows: a visual programming platform for text mining and natural language processing", *Sci. comput. program.*, vol. 121, pp. 128-152, 2016.
24. Tina Ravnikar, Marko Bohanec, Gregor Muri, "Monitoring and assessment of anthropogenic activities in mountain lakes: a case of the Fifth Triglav Lake in the Julian Alps", *Environ. monit. assess.*, vol. 188, pp. 253-1-253-17, mar. 2016.
25. Yves Scherrer, Tomaž Erjavec, "Modernising historical Slovene words", *Nat. lang. eng.*, vol. 22, no. 6, pp. 881-905, 2016.
26. Nikola Simidjievski, Ljupčo Todorovski, Sašo Džeroski, "Modeling dynamic systems with efficient ensembles of process-based models", *PloS one*, vol. 11, no. 4, pp. 0153507-1-0153507-27, 2016.
27. Ana Marija Sobočan, Senja Pollak, "Manifestacije moralne panike v medijskih podobah novele družinske zakonodaje", *Čas. krit. znan.*, vol. 44, no. 266, pp. 163-187, 2016.
28. Darja Šemrov, Rok Marsetič, Marijan Žura, Ljupčo Todorovski, Aleksander Srdič, "Reinforcement learning approach for train rescheduling on a single-track railway", *Transp. res., Part B: methodol.*, vol. 86, no. apr., pp. 250-267, 2016.
29. Bosiljka Tadić, Miroslav Andjelković, Biljana Mileva-Boshkoska, Zoran Levnajić, "Algebraic topology of multi-brain connectivity networks reveals dissimilarity in functional patterns during spoken communications", *PloS one*, vol. 11, no. 11, pp. e0166787-1-e0166787-25, 2016.
30. Jovan Tanevski, Ljupčo Todorovski, Sašo Džeroski, "Learning stochastic process-based models of dynamical systems from knowledge and data", *BMC systems biology*, vol. 10, pp. 30-1-30-17, 2016.
31. Jovan Tanevski, Ljupčo Todorovski, Sašo Džeroski, "Process-based design of dynamical biological systems", *Scientific reports*, vol. 6, pp. 34107-1-34107-13, 2016.
32. Vedrana Vidulin, Tomislav Šmuc, Fran Supek, "Extensive complementarity between gene function prediction methods", *Bioinformatics (Oxf.)*, vol. 32, no. 23, pp. 3645-3653, 2016.

Published Conference Contribution

1. Darko Aleksovski, Dejan Dovžan, Sašo Džeroski, Juš Kocijan, "A comparison of fuzzy identification methods on benchmark datasets", In: *Proceedings of the 4th IFAC Conference on Intelligent Control and Automation, ICONS 2016, 1-3 June 2016, Reims, France*, (IFAC papersOnline, Vol. 49, no. 5, 2016), pp. 31-36.
2. Špela Arhar Holdt, Darja Fišer, Tomaž Erjavec, Simon Krek, "Syntactic annotation of Slovene CMC: first steps", In: *Proceedings of the 4th Conference on CMC and Social Media Corpora for the Humanities, 27-28 September 2016, Faculty of Arts, University of Ljubljana, Ljubljana, Slovenia*, pp. 3-6.
3. Renato Barišić, Biljana Mileva-Boshkoska, "MIMOSA OSA-EAI: an information exchange standard: standard razmjene informacija", In: *Kvaliteta čini razliku: zbornik radova: proceedings*, pp. 45-57.
4. Michael Beißwenger *et al.* (12 authors), "Integrating corpora of computer-mediated communication into the language resources landscape: initiatives and best practices from French, German, Italian and Slovenian projects", In: *Proceedings*, pp. 1-5.
5. Jože Bučar, Janez Povh, Martin Žnidaršič, "Sentiment classification of the Slovenian news texts", In: *Proceedings of the 9th International Conference on Computer Recognition Systems CORES 2015*, (Advances in intelligent systems and computing, Vol. 403, 2016), pp. 777-787.
6. Bojan Cestnik, Elsa Fabbretti, Donatella Gubiani, Nada Lavrač, Tanja Urbančič, "Exploring outlier documents to investigate potential links between gut microbiota and Alzheimer's disease", In: *Proceedings, extended abstracs*, IWBBIO 2016, International Work-Conference on Bioinformatics and Biomedical Engineering, April 20-22, 2016, Granada, Spain, pp. 475-486.
7. Bojan Cestnik, Natalia Manakova, Aram Karalič, Tanja Urbančič, "Engaging the net generation students in enhanced teaching and leaning with QTvity: the next steps", In: *Proceedings: European thematic network, future education and training in computing: how to support learning at anytime anywhere*, International Conference on e-Learning, 8-9 September, Bratislava, Slovakia, pp. 86-93.
8. Jaka Čibej, Špela Arhar Holdt, Tomaž Erjavec, Darja Fišer, "Razvoj učne množice za izboljšano označevanje spletnih besedil", In: *Zbornik konference Jezikovne tehnologije in digitalna humanistika, 29. september - 1. oktober 2016, Filozofska fakulteta, Univerza v Ljubljani, Ljubljana, Slovenija*, pp. 40-46.
9. Jaka Čibej, Darja Fišer, Tomaž Erjavec, "Normalisation, tokenisation and sentence segmentation of Slovene tweets", In: *Normalisation and analysis of social media texts (NormSoMe): [workshop proceedings]*, pp. 5-10.
10. Marko Debeljak, Andrej Ficko, Robert Brus, "Ocena ogroženosti avtohtonih drevesnih vrst pred invazivnimi tujerodnimi vrstami na primeru črnega topola (Populus nigra L.): the case of Black poplar (Populus nigra L.)", In: *Invazivne tujerodne vrste v gozdovih ter njihov vpliv na trajnostno rabo gozdnih virov: zbornik prispevkov posvetovanja z mednarodno udeležbo: lectures presented at the conference with international participation*, 2016, pp. 119-124.
11. Tomaž Erjavec, Špela Arhar Holdt, Jaka Čibej, Kaja Dobrovoljc, Darja Fišer, Cyprian Laskowski, Katja Zupan, "Annotating CLARIN.SI TEI corpora with WebAnno", In: *Proceedings*, 2016, pp. 1-5.
12. Tomaž Erjavec, Jaka Čibej, Špela Arhar Holdt, Nikola Ljubešić, Darja Fišer, "Gold-standard datasets for annotation of Slovene computer-mediated communication", In: *RASLAN 2016: Recent Advances in Slavonic Natural Language Processing: proceedings*, pp. 29-40.
13. Tomaž Erjavec, Darja Fišer, Nikola Ljubešić, Nataša Logar, Milan Ojsteršek, "Slovenska znanstvena besedila: prototipni korpus in načrt analiz", In: *Zbornik konference Jezikovne tehnologije in digitalna humanistika, 29. september - 1. oktober 2016, Filozofska fakulteta, Univerza v Ljubljani, Ljubljana, Slovenija*, pp. 58-64.
14. Darja Fišer, Tomaž Erjavec, "Analysis of sentiment labeling of Slovene user-generated content", In: *Proceedings of the 4th Conference on CMC and Social Media Corpora for the Humanities, 27-28 September 2016, Faculty of Arts, University of Ljubljana, Ljubljana, Slovenia*, pp. 22-25.
15. Darja Fišer, Nikola Ljubešić, "Detecting semantic shifts in Slovene twitterese", In: *RASLAN 2016: Recent Advances in Slavonic Natural Language Processing: proceedings*, pp. 43-50.
16. Darja Fišer, Jasmina Smailović, Tomaž Erjavec, Igor Mozetič, Miha Grčar, "Sentiment annotation of Slovene user-generated content", In: *Zbornik konference Jezikovne tehnologije in digitalna humanistika, 29. september - 1. oktober 2016, Filozofska fakulteta, Univerza v Ljubljani, Ljubljana, Slovenija*, pp. 65-70.
17. Darja Fišer, Vit Suchomel, Miloš Jakubiček, "Terminology extraction for academic Slovene using sketch engine", In: *RASLAN 2016: Recent Advances in Slavonic Natural Language Processing: proceedings*, pp. 135-141.
18. Dragan Gamberger, Bernard Ženko, Alexis Mitelpunkt, Nada Lavrač, "Homogeneous clusters of Alzheimer's disease patient population", In: *Selected articles of the 3rd International Work-Conference on*

Bioinformatics and Biomedical Engineering (IWBBIO 2015), 15-17 April 2015, Grenada, Spain, (BioMedical engineering online, Vol. 15, suppl. 1, 2016), pp. 578-1-578-34.
19. Polona Gantar, Iza Škrjanec, Darja Fišer, Tomaž Erjavec, "Slovar tviterščine", In: *Zbornik konference Jezikovne tehnologije in digitalna humanistika, 29. september - 1. oktober 2016, Filozofska fakulteta, Univerza v Ljubljani, Ljubljana, Slovenija*, pp. 71-76.
20. Dimitros Gatsios, George Rigas, Dragana Miljković, Barbara Koroušić-Seljak, Marko Bohanec, María Soledad Arredondo, Angelo Antonini, Spyros Konitsiotis, Dimitrios I. Fotiadis, "Mhealth platform for Parkinson's disease management", In: *Final program*, CBHI 2016, 18th International Conference on Biomedicine and Health Informatics, Februar 25-26, Dubai, UAE, 5 pp..
21. Teja Goli, Eneja Osrajnik, Darja Fišer, "Analiza krajšanja slovenskih sporočil na družbenem omrežju Twitter", In: *Zbornik konference Jezikovne tehnologije in digitalna humanistika, 29. september - 1. oktober 2016, Filozofska fakulteta, Univerza v Ljubljani, Ljubljana, Slovenija*, pp. 77-82.
22. Jan Kralj, Marko Robnik Šikonja, Nada Lavrač, "Heterogeneous network decomposition and weighting with text mining heuristics", In: *New frontiers in mining complex patterns: 4th International Workshop, NFMCP 2015 held in conjunction with ECML-PKDD 2015 Porto, Portugal, September 7, 2015: revised selected papers*, (Lecture notes in computer science, Lecture notes in artificial intelligence, 9607), pp. 194-208.
23. Jan Kralj, Anže Vavpetič, Michel Dumontier, Nada Lavrač, "Network ranking assisted semantic data mining", In: *Bioinformatics and Biomedical Engineering, IWBBIO 2016*, (Lecture notes in computer science, 9656), pp. 752-764.
24. Vladimir Kuzmanovski, Aneta Trajanov, Sašo Džeroski, Marko Debeljak, "Integrating predictive and decision modelling in decision support system for water protection from phytochemicals", In: *Zbornik*, 8. študentska konferenca Mednarodne podiplomske šole Jožefa Stefana, 31. maj in 1. junij 2016, Ljubljana, Slovenija, pp.165-174.
25. Agnieszka Lawrynowicz, Diego Esteves, Panče Panov, Tommaso Soru, Sašo Džeroski, Joaquin Vanschoren, "An algorithm, implementation and execution ontology design pattern", In: *WOP 2016: Worskop on Ontology and Semantic Web Patterns (7th edition) 2016, 18th October 2016, Kobe, Japan*, 12 pp..
26. Nikola Ljubešić, Tomaž Erjavec, "Corpus vs. lexicon supervision in morphosyntactic tagging: the case of Slovene", In: *LREC 2016: [proceedings]*, Tenth International Conference on Language Resources and Evaluation, Portorož, Slovenia 23-28 May, 2016, pp. 1527-1531.
27. Nikola Ljubešić, Tomaž Erjavec, Darja Fišer, "Corpus-based diacritic restoration for South Slavic languages", In: *LREC 2016: [proceedings]*, Tenth International Conference on Language Resources and Evaluation, Portorož, Slovenia 23-28 May, 2016, pp. 3612-3616.
28. Nikola Ljubešić, Tomaž Erjavec, Darja Fišer, Tanja Samardžić, Maja Miličević, Filip Klubička, Filip Petkovski, "Easily accessible language technologies for Slovene, Croatian and Serbian", In: *Zbornik konference Jezikovne tehnologije in digitalna humanistika, 29. september - 1. oktober 2016, Filozofska fakulteta, Univerza v Ljubljani, Ljubljana, Slovenija*, pp. 120-124.
29. Nikola Ljubešić, Darja Fišer, "A global analysis of emoji usage", In: *Proceedings of the 10th Web as Corpus Workshop (WAC-X) and The EmpiriST Shared Task: ACL 2016*, The 10th Web as Corpus Workshop (WAC-X) and The EmpiriST Shared Task, August 12, Berlin, Germany, pp. 82-89.
30. Nikola Ljubešić, Darja Fišer, "Private or Corporate? Predicting user types on Twitter", In: *WNUT 2016: proceedings of the Workshop*, 2016, pp. 4-12.
31. Nikola Ljubešić, Darja Fišer, "Slovene Twitter analytics", In: *Proceedings of the 4th Conference on CMC and Social Media Corpora for the Humanities, 27-28 September 2016, Faculty of Arts, University of Ljubljana, Ljubljana, Slovenia*, pp. 39-43.
32. Nikola Ljubešić, Tanja Samardžić, Curdin Derungs, "TweetGeo - a tool for collecting, processing and analysing geo-encoded linguistic data", V: *Proceedings of COLING 2016: technical papers*, The 26th International Conference on Computational Linguistics, December 11-16, 2016, Osaka, Japan, str. 3412-3421, 2016.
33. Nikola Ljubešić, Katja Zupan, Darja Fišer, Tomaž Erjavec, "Normalising Slovene data: historical texts vs. user-generated content", In: *Proceedings of the 13th Conference on Natural Language Processing (KONVENS), September 19-21, 2016, Bochum, Germany*, (Bochumer linguistische Arbeitsberichte, 16), pp. 146-155.
34. Mitja Luštrek, Erik Dovgan, Aljoša Vodopija, Marko Bohanec, Anneleen Baert, Sofie Pardaens, Els Clays, "Aplikacija za pomoč pri telesni vadbi bolnikov s srčnim popuščanjem", In: *Delavnica Elektronsko in mobilno zdravje: zbornik 19. mednarodne multikonference Informacijska družba - IS 2016, 10.-11. oktober 2016, [Ljubljana, Slovenija]*, pp. 61-62.
35. Pedro Martins, Senja Pollak, Tanja Urbančič, Amilcar Cardoso, "Optimality principles in computational approaches to conceptual blending: do we need them (at) all?", In: *Proceedings of the 7th International Conference on Computational Creativity, ICCC 2016, 27 June - 1 July*, pp. 346-353.
36. Matej Mihelčić, Sašo Džeroski, Nada Lavrač, Tomislav Šmuc, "Redescription mining with multi-target predictive clustering trees", In: *New frontiers in mining complex patterns: 4th International Workshop, NFMCP 2015 held in conjunction with ECML-PKDD 2015 Porto, Portugal, September 7, 2015: revised selected papers*, (Lecture notes in computer science, Lecture notes in artificial intelligene, 9607), pp. 125-143, 2016.
37. Vanja Mileski, Dragi Kocev, Bogdan Draganski, Sašo Džeroski, "Multi-dimensional analysis of PPMI data", In: *Zbornik*, 8. študentska konferenca Mednarodne podiplomske šole Jožefa Stefana, 31. maj in 1. junij 2016, Ljubljana, Slovenija, pp. 175-179.
38. Biljana Mileva-Boshkoska, Marko Bohanec, Theo W. Prins, Esther Kok, "Usage of SIGMO, a decision support system for the assessment of genetically modified organisms in food and feed", In: *Slovenska konferenca o umetni inteligenci: zbornik 19. mednarodne multikonference Informacijska družba - IS 2016, 12. oktober 2016, [Ljubljana, Slovenija]: zvezek A*, pp. 5-8.
39. Dragana Miljković, Nada Lavrač, Marko Bohanec, Biljana Mileva-Boshkoska, "Detection of dependencies between literature domains through relation extraction and copulas", In: *International Conference on Control, Decision and Information Technologies, CoDIt 2016, April 6-8, 2016, St. Paul's Bay, Malta*, pp. 302-307.
40. Aljaž Osojnik, Sašo Džeroski, Dragi Kocev, "Option predictive clustering trees for multi-target regression", In: *Discovery science: 19th International Conference, DS 2016 Bari, Italy, October 19-21, 2016: proceedings*, (Lecture notes in computer science, Lecture notes in artificial intelligence, LNCS 9956), pp. 118-133, 2016.
41. Aljaž Osojnik, Panče Panov, Sašo Džeroski, "Comparison of tree-based methods for multi-target regression on data streams", In: *New frontiers in mining complex patterns: 4th International Workshop, NFMCP 2015 held in conjunction with ECML-PKDD 2015 Porto, Portugal, September 7, 2015: revised selected papers*, (Lecture notes in computer science, Lecture notes in artificial intelligene, 9607), pp. 17-31, 2016.
42. Matej Petković, Panče Panov, Sašo Džeroski, "A comparison of different data transformation approaches in the feature ranking context", In: *Discovery science: 19th International Conference, DS 2016 Bari, Italy, October 19-21, 2016: proceedings*, (Lecture notes in computer science, Lecture notes in artificial intelligence, LNCS 9956), pp. 310-324, 2016.
43. Senja Pollak, Biljana Mileva-Boshkoska, Dragana Miljković, Geraint Wiggins, Nada Lavrač, "Computational creativity conceptualisation grounded on ICCC papers", In: *Proceedings of the 7th International Conference on Computational Creativity, ICCC 2016, 27 June - 1 July*, pp. 123-130.
44. Damjan Popič, Darja Fišer, Katja Zupan, Polona Logar, "Raba vejice v uporabniških spletnih vsebinah", In: *Zbornik konference Jezikovne tehnologije in digitalna humanistika, 29. september - 1. oktober 2016, Filozofska fakulteta, Univerza v Ljubljani, Ljubljana, Slovenija*, pp. 106-110.
45. Aleksandra Rashkovska, Dragi Kocev, Roman Trobec, "Clustering of heartbeats from ecg recordings obtained with wireless body sensors", In: *Mipro proceedings*, (MIPRO ... (Tisak)), MIPRO 2016, 39th International Convention, May 30-June 3, 2016, Opatija, Croatia, pp. 481-486.
46. Nikola Simidjievski, Ljupčo Todorovski, Sašo Džeroski, "Learning ensembles of process-based models by bagging of random library samples", In: *Discovery science: 19th International Conference, DS 2016 Bari, Italy, October 19-21, 2016: proceedings*, (Lecture notes in computer science, Lecture notes in artificial intelligence, LNCS 9956), pp. 245-260, 2016.
47. Borut Sluban, Miha Grčar, Igor Mozetič, "Temporal multi-layer network construction from major news events", In: *Complex networks VII: proceedings of the 7th Workshop on Complex Networks, CompleNet 2016, March 23 to 25, 2016, Bourgogne, Dijon, France*, (Studies in computational intelligence, volume 644), pp. 29-41.
48. Borut Sluban, Jasmina Smailović, Igor Mozetič, "Understanding financial news with multi-layer network analysis", In: *Proceedings of ECCS 2014: European Conference on Complex Systems, 22 to 26 September 2014, Lucca, Italy*, 2016, pp. 193-207.
49. Jasmina Smailović, Miha Grčar, Nada Lavrač, Martin Žnidaršič, "Dynamic parameter adaptation of SVM based active learning methodology", In: *AL@iKNOW 2016: proceedings of the Workshop on Active Learning: Applications, Foundations and Emerging Trends co-*

located with International Conference on Knowledge Technologies and Data-driven Business (i-KNOW 2016), Graz Austria, October 18, 2016, (CEUR workshop proceedings, vol. 1707).

50. Jasmina Smailović, Martin Žnidaršič, Nada Lavrač, Senja Pollak, Janez Kranjc, "Modelling human appreciation of machine generated What-if ideas", In: *Constructive machine learning: NIPS 2016, Neural Information Processing Systems, December 10, 2016, Barcelona, Spain*, 5 pp.
51. Larisa N. Soldatova, Panče Panov, Sašo Džeroski, "Ontology engineering: from an art to a craft", In: *Ontology engineering: 12th International Experiences and Directions Workshop on OW, OWLED, 2015, Co-located with ISWC 2015, Bethlehem, Pa, USA, October 9-10, 2015: revised selected papers*, (Lecture notes in computer science, vol. 9557), pp. 174-181, 2016.
52. Tomaž Stepišnik Perdih, Panče Panov, Andrej Bauer, Sašo Džeroski, "Primerjava razdalj pri klasifikaciji in regresiji", In: *Slovenska konferenca o umetni inteligenci: zbornik 19. mednarodne multikonference Informacijska družba - IS 2016, 12. oktober 2016, [Ljubljana, Slovenija]: zvezek A*, pp. 68-71.
53. Iza Škrjanec, Senja Pollak, "Topic ontologies of the Slovene blogosphere: a gender perspective", In: *Proceedings of the 4th Conference on CMC and Social Media Corpora for the Humanities, 27-28 September 2016, Faculty of Arts, University of Ljubljana, Ljubljana, Slovenia*, pp. 62-65.
54. Tadej Škvorc, Nada Lavrač, Marko Robnik Šikonja, "Co-bidding graphs for constrained paper clustering", In: *5th Symposium on Languages, Applications and Technologies: SLATE'16, June 20-21, 2016, Maribor, Slovenia*, (Open access series in informatics, vol. 51), pp. 1-13.
55. Alan Tapscott, Javier Gomez, Pablo Gervas, Jasmina Smailović, Martin Žnidaršič, "Empirical evidence of the limits of automatic assessment of fictional ideation", In: *Proceedings of the Workshop on Computational Creativity, Concept Invention, and General Intelligence (C3GI 2016) co-located with the 28th European Summer School in Logic, Language and Information (ESSLLI 2016), Bozen-Bolzano, Italy, August 20-22, 2016*, 14 pp..
56. Anita Valmarska, Dragana Miljković, Nada Lavrač, Marko Robnik Šikonja, "Towards multi-view approach to Parkinson's disease quality of life data analysis", In: *ECML-PKDD 2016, The European Conference on Machine Learning and Principles and Practice of Databases: 5th International Workshop on New Frontiers in Mining Complex Patterns, 19-23 September 2016, 2016, Riva de Garda, Italy*, 12 pp.
57. Vedrana Vidulin, Marija Brbić, Fran Supek, Tomislav Šmuc, "Evaluation of fusion approaches in large-scale bio-annotation setting", In: *4th Workshop on Machine Learning in Life Science, 23 September 2016, Riva del Garda, Italy*, pp. 37-51.
58. Darko Zelenika, Janez Povh, Bernard Ženko, "Text detection in document images by machine learning algorithms", In: *Proceedings of the 9th International Conference on Computer Recognition Systems CORES 2015*, (Advances in intelligent systems and computing, Vol. 403), pp. 169-179.
59. Katja Zupan, Tomaž Erjavec, "Generiranje kritičnih prepisov s strojnim prevajanjem na ravni znakov", In: *Zbornik konference Jezikovne tehnologije in digitalna humanistika, 29. september - 1. oktober 2016, Filozofska fakulteta, Univerza v Ljubljani, Ljubljana, Slovenija*, pp. 175-179.
60. Ana Zwitter Vitez, Darja Fišer, "Linguistic analysis of emotions in online news comments: an example of the Eurovision song contest", In: *Proceedings of the 4th Conference on CMC and Social Media Corpora for the Humanities, 27-28 September 2016, Faculty of Arts, University of Ljubljana, Ljubljana, Slovenia*, pp. 74-76.
61. Martin Žnidaršič, Janez Kranjc, Nada Lavrač, "Computational creativity infrastructure for online software composition: a conceptual blending use case", In: *Proceedings of the 7th International Conference on Computational Creativity, ICCC 2016, 27 June - 1 July*, pp. 371-378.

Independent Component Part or a Chapter in a Monograph

1. Tina Jukić, Ljupčo Todorovski, András Nemeslaki, "Iskanje identitete raziskovalnega področja e-uprave", In: *Slovenska javna uprava na razpotju strateških sprememb*, (Upravna misel), Aleksander Aristovnik, ed., et al, 1. natis, Ljubljana, Fakulteta za upravo, 2016, pp. 111-136.
2. Petra Kralj Novak, Nada Lavrač, Geoffrey I. Webb, "Supervised descriptive rule induction", In: *Encyclopedia of machine learning and data mining*, Claude Sammut, ed., Geoffrey I. Webb, ed., Heidelberg [etc.], Springer, 2016, 4 pp..
3. Jan Kralj, Anita Valmarska, Miha Grčar, Marko Robnik Šikonja, Nada Lavrač, "Analysis of text-enriched heterogeneous information networks", In: *Big data analysis: new algorithms for a new society*, (Studies in big data, volume 16), Nathalie Japkowicz, ed., Jerzy Stefanowski, ed., Cham, Springer, 2016, pp. 115-139.
4. Nataša Logar, Špela Arhar Holdt, Tomaž Erjavec, "Slovenski strokovni jezik: korpusni opis trpnika", In: *Toporišičeva obdobja*, (Obdobja, Simpozij, = Symposium, 35), Erika Kržišnik, ed., Miran Hladnik, ed., 1. natis, Ljubljana, Znanstvena založba Filozofske fakultete, 2016, pp. 237-245.
5. Dragana Miljković, Darko Aleksovski, Vid Podpečan, Nada Lavrač, Bernd Malle, Andreas Holzinger, "Machine learning and data mining methods for managing Parkinson's disease", In: *Machine learning for health informatics: state-of-the-art and future challenges*, (Lecture notes in computer science, Lecture notes in artificial intelligence, LNAI 9605), Andreas Holzinger, ed., Cham [etc.], Springer, 2016, lNAI 9605, pp. 209-220, 2016.
6. Matic Perovšek, Matjaž Juršič, Bojan Cestnik, Nada Lavrač, "Empowering bridging term discovery for cross-domain literature mining in the textflows platform", In: *Machine learning for health informatics: state-of-the-art and future challenges*, (Lecture notes in computer science, Lecture notes in artificial intelligence, LNAI 9605), Andreas Holzinger, ed., Cham [etc.], Springer, 2016, lNAI 9605, pp. 59-98, 2016.
7. Erich Prunč, Matija Ogrin, Tomaž Erjavec, "Kapelski pasijon v elektronski znanstvenokritični izdaji. "Nov" rokopis, nove raziskovalne poti", In: *Toporišičeva obdobja*, (Obdobja, Simpozij, = Symposium, 35), Erika Kržišnik, ed., Miran Hladnik, ed., 1. natis, Ljubljana, Znanstvena založba Filozofske fakultete, 2016, pp. 395-402.
8. Tanja Urbančič, Davor Orlič, "Opening up education: towards affordable and sustainable solutions", In: *Developing successful strategies for global policies and cyber transparency in E-learning*, Gulsun Eby, ed., T. Volkan Yuzer, ed., Simber Atay, ed., Hershey, Information Science Reference, 2016, pp. 133-144.
9. Špela Vintar, Darja Fišer, "Using WordNet-based word sense disambiguation to improve MT performance", In: *Hybrid approaches to machine translation*, (Theory and applications of natural language processing), Marta Ruiz, ed., Berlin, Springer, 2016, pp. 191-205.

Encyclopaedia, Lexicon, Dictionary Entry

1. Kristina Hmeljak Sangawa, Tomaž Erjavec, Irena Srdanović, *Japanese-Slovene learner's dictionary jaSlo 3.1: Slovenian language resource repository CLARIN.SI*, Ljubljana, Department of Asian Studies, [Faculty of Arts], University of Ljubljana, 2016.
2. Tomaž Erjavec, *Lexicon of historical Slovene imp25k 1.1*, Ljubljana, Slovenian Language Resource Repository CLARIN.SI, 2014.
3. Tomaž Erjavec *et al.* (14 authors), *MULTEXT-East non-commercial lexicons 4.0*, Ljubljana, Slovenian Language Resource Repository CLARIN.SI, 2010.
4. Tomaž Erjavec *et al.* (19 authors), *MULTEXT-East free lexicons 4.0*, Ljubljana, Slovenian Language Resource Repository CLARIN.SI, 2010.

Mentoring

1. Vladimir Kuzmanovski, *Integrating decision support and data mining for risk evaluation and management: a methodological framework and a case study in agriculture:* doctoral dissertation, Ljubljana, 2016 (mentor Marko Debeljak; co-mentor Sašo Džeroski).
2. Matic Perovšek, *Advanced text processing workflows in a web-based text mining platform:* doctoral dissertation, Ljubljana, 2016 (mentor Bojan Cestnik; co-mentor Nada Lavrač).
3. Nikola Simidjievski, *Ensembles of process-based models of dynamic systems:* doctoral dissertation, Ljubljana, 2016 (mentor Sašo Džeroski; co-mentor Ljupčo Todorovski).
4. Jovan Tanevski, *Deterministic and stochastic process-based modeling and design of dynamic systems in biology:* doctoral dissertation, Ljubljana, 2016 (mentor Sašo Džeroski; co-mentor Ljupčo Todorovski).
5. Anže Vavpetič, *Semantic subgroup discovery:* doctoral dissertation, Ljubljana, 2016 (mentor Nada Lavrač).
6. Darja Šemrov, *Railway traffic scheduling with use of reinforcement learning:* doctoral dissertation, Ljubljana, 2016 (mentor Marijan Žura; co-mentor Ljupčo Todorovski).
7. Katarina Klanšek, *Analiza vpliva zunanjega izvajanja IKT storitev na konkurenčnost slovenskih podjetij:* master's thesis, Ljubljana, 2016 (mentor Bojan Cestnik).
8. Anamarija Leben, *Life-event modelling for development of active e-government portals:* master's thesis, Ljubljana, 2016 (mentor Ljupčo Todorovski).

9. Primož Njegač, *Implementacija sistema za upravljanje odnosov s strankami (CRM) v podjetje trgovina Njegač, d.o.o.:* master's thesis, Ljubljana, 2016 (mentor Bojan Cestnik).
10. Janez Pagon, *Floodplain forests of Soča river between Kobarid and confluence with river Tolminka: current situation and development:* master's thesis, Nova Gorica, 2016 (mentor Marko Debeljak).
11. Anja Šostar, *Increasing customer loyalty with data mining methods:* master's thesis, Ljubljana, 2016 (mentor Irena Ograjenšek; co-mentor Nada Lavrač).
12. Jože Železnikar, *Analysis of digital signing in the cloud:* master's thesis, Ljubljana, 2016 (mentor Ljupčo Todorovski).
13. Jerneja Boslič, *Analysis of the user satisfaction with the government information system for European Union affairs:* master's thesis, Ljubljana, 2016 (mentor Ljupčo Todorovski).
14. Blaž Bunc, *Analysis of the dynamic of copyright infringement around the world:* master's thesis, Ljubljana, 2016 (mentor Ljupčo Todorovski).
15. Gašper Cesar, *Comparative analysis of internet neutrality in chosen countries:* master's thesis, Ljubljana, 2016 (mentor Ljupčo Todorovski).
16. Danjel Cotič, *Improving the system of monitoring employees competences in Mahle Letrika company:* master's thesis, Nova Gorica, 2016 (mentor Tanja Urbančič).
17. Klaudija Habjan, *The analysis of students' characteristics related to the achievement in the introduction course to programming:* master's thesis, Ljubljana, 2016 (mentor Irena Nančovska Šerbec; co-mentor Ljupčo Todorovski).
18. Dušan Likar, *Študija vključitve zunanjih dobaviteljev v enotni poslovni informacijski sistem podjetja Pipistrel:* master's thesis, Nova Gorica, 2016 (mentor Bojan Cestnik).
19. Senada Medara, *Comparing the development of information society in Slovenia with selected EU countries:* master's thesis, Ljubljana, 2016 (mentor Ljupčo Todorovski).
20. Primož Pahor, *Assessment and selection of technical solutions for stator production:* master's thesis, Nova Gorica, 2016 (mentor Marko Bohanec).
21. David Rejc, *Selecting soldering technology: A comparison of multi-attribute decision-making methods:* master's thesis, Nova Gorica, 2016 (mentor Marko Bohanec).
22. Tomaž Stepišnik Perdih, *Razdalje na strukturiranih podatkih:* master's thesis, Ljubljana, 2016 (mentor Sašo Džeroski; co-mentor Andrej Bauer).
23. Jernej Ušaj, *Selecting a hydraulic wood splitter using multi-criteria assessment and financial analysis:* master's thesis, Nova Gorica, 2016 (mentor Marko Bohanec).
24. Johannes Vuga Gregorič, *A decision support system for selecting forms of online advertising messages to inform selected categories of users:* master's thesis, Nova Gorica, 2016 (mentor Marko Bohanec).