

DEPARTMENT OF KNOWLEDGE TECHNOLOGIES

E-8

The Department of Knowledge Technologies performs research in advanced information technologies aimed at acquiring, storing and managing knowledge to be used in the development of knowledge-based applications. Established areas include intelligent data analysis (machine learning, data mining, and knowledge discovery in databases), semantic data mining and the semantic web, language technologies and computational linguistics, decision support and knowledge management. Apart from research in knowledge technologies, we are also developing applications in environmental sciences and ecology, medicine and health care, biomedicine and bioinformatics, economy and marketing.

In 2012 we were involved in eight national and thirteen European projects, most of them funded in the seventh framework program (FP7). In terms of our collaboration in EU projects, we were the most successful program group in Slovenia.

In the area of intelligent data analysis and data mining we have developed several new methods and used them in a number of application areas. We developed a new approach for creating and implementing data-mining workflows, upgrading the service-oriented knowledge-discovery platform Orange4WS, also developing a new workflow platform called ClowdFlows. In Orange4WS we also developed and implemented a novel SDM-Toolkit used for semantic subgroup discovery; a new triplet extraction methodology from text, used for signaling network construction in the modeling of plant defense response; and the SegMine methodology, which together with the BioMine system for the detection of new links between genes, allows for semantic gene expression analysis by using bio-ontologies as background knowledge.

We developed a new system for bisociative knowledge discovery called CrossBee, used for discovering new connections between different medical domains. We developed a new methodology for mining document-enriched heterogeneous information networks and a new approach to equation discovery, enabling the learning of process-based models of dynamic systems from data and domain knowledge, which can use different criteria and methods for parameter optimization. We developed new methods for learning rule ensembles for multi-target regression and new methods for multi-label classification and conducted a large-scale empirical comparison of existing ones. We used the newly developed methods in a number of environmental problems, including the hierarchical classification of diatoms from images, modeling the dynamics of aquatic ecosystems, and modeling gene flow between conventional and genetically modified crops.

We have successfully concluded our work on the FP7 project on data mining called e-LICO (e-Laboratory for Interdisciplinary Collaborative Research in Data Mining and Data-Intensive Sciences) where we developed new web services for subgroup discovery.

We successfully concluded our collaboration within the FP7 project PHAGOSYS (Systems biology of phagosome formation and maturation – modulation by intracellular pathogens), where we used equation-discovery approaches to learn models of the dynamics of endocytosis, an important process of immune response, and analyzed high throughput data from compound and genomic screens about different aspects of the same process (e.g., in response to different pathogens). We continued our collaboration in two other FP7 projects: SUMO, where we are developing methods for learning supermodels (ensemble models of dynamic systems), and REWIRE, where we apply machine-learning methods to analyze patient-record and sensor data in the context of post-stroke rehabilitation. In September we started work on the EU 7FP FET project from the field of computer understanding of natural language called MUSE (Machine Understanding for interactive StoryTElling), the goal of which is to illustrate texts as 3D animation. In the area of text and web mining and heterogeneous information network analysis we continued our research in the framework of three European FP7 projects: FIRST (Large scale information extraction and integration infrastructure for supporting financial decision making), FOC (Forecasting Financial Crises), and ENVISION (ENVironmental Services Infrastructure with ONtologies). In FIRST and FOC we focus on analyzing large amounts of dynamic and heterogeneous sources of financial information and develop online data-mining



Head:

Prof. Nada Lavrač

We produced the first digital dictionary of historical Slovene, containing over 20,000 entries and available from <http://nl.ijs.si/imp/>, which can be used by linguists, as well as for applications in human language technologies, e.g., to support a full-text search in digital libraries of Slovenian cultural heritage.

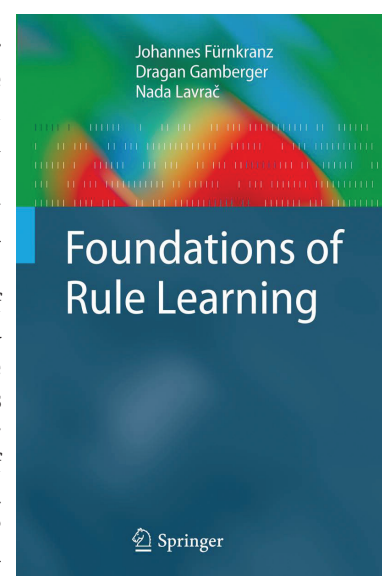


Figure 1: Monograph "Foundations of Rule Learning" (Springer 2012, 334 pages), co-authored by Prof. Nada Lavrač, presents the foundations, techniques and selected applications of rule learning as investigated in classic machine learning and modern data mining.

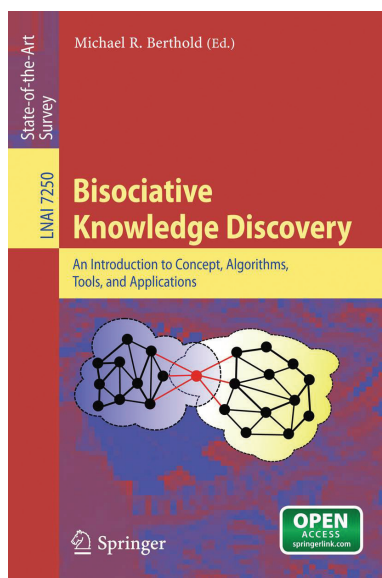


Figure 2: Result of successful EU FP7 project BISON is an open-access book "Bisociative Knowledge Discovery", Springer 2012. Members of the Department of Knowledge Technologies at JSI contributed nine book chapters addressing methods for discovering new, domain-bridging connections and patterns from heterogeneous data sources.

tools for near-real-time processing of vast amounts of constantly evolving data (financial news, blogs and tweets). We permanently monitor 200 financial websites and capture about 40,000 documents per day. In 2012 we focused on data analysis and end-user solutions, such as sentiment analysis on financial products, estimating the reputation of financial institutions, and on online fraud detection. In FOC we extract indicators based on the sentiment analysis of large streams of textual data with the goal of forecasting financial crises. In ENVISION we develop tools for multilingual support for ontology management and a semantic description of geographic data and services. The tools are integrated into a web application and constitute one of the modules of the on-line environmental decision support portal.

During the 2012 Slovenian presidential elections, we launched a novel public sentiment monitoring solution in collaboration with Gama System and POP TV. We developed a sentiment analysis prototype that collected and analyzed tweets about the three presidential candidates. The sentiment charts were shown on prime time during live TV debates on POP TV and were controversial in the sense that they were in conflict with the polls carried out by the major polling agencies in Slovenia. The polling agencies predicted the then-current president would win the first round of elections, while our system clearly showed the lead of Borut Pahor, who in the end won both the first and the final round of elections.

In the area of language technologies we continued work on the development of language resources and technologies for historical Slovene, where we produced a manually annotated corpus (300,000 words), a digital lexicon (20,000 entries) and an annotation tool for historical Slovene, with which we then annotated a large collection (400 books, 30,000 pages) of historical Slovene texts and made it available as a digital library and through a web-based concordancer. With this we are enabling empirically based studies of diachronic Slovene, as well as better accessibility of Slovene written cultural heritage, already used in the dLib.si digital library portal of the National and University Library (NUL) of Slovenia. This work was undertaken in the scope of the FP7 project IMPACT (Improving Access to Text) in which we collaborate with NUL and as part of the Google awarded project "Developing language models for historical Slovene", in which we collaborate with the Scientific Research Centre of the Slovene Academy of Sciences and Arts (SRC SASA). In the scope of the national project "Unknown 17th and 18th century manuscripts of Slovene literature: information-technology aided register, scholarly editions and analyses" we finished our joint work with SRC SASA, where we implemented a Fedora Commons platform for searching manuscript descriptions and viewing manuscripts. We also started joint work with SRC SASA on the national project "Leading Slovene Humanists from the 16th to the mid-19th Century".

In the scope of the project "Communication in Slovene", under the lead of the company Amebis we finished work on a new generation of reference corpora for Slovene (e.g., Gigafida with more than a billion words), which also include large freely available corpora, so far lacking for the Slovene language. We completed our work on the linguistic annotation of parallel bi-lingual corpora, performed in the scope of the national project "Slovene translation studies – resources and research" under the lead of the Department for Translation Studies at the Arts Faculty of the University of Ljubljana. The corpora are used for linguistic studies of translation processes, while also being useful for the development of multilingual language technologies. In cooperation with the same department we are continuing work on enlarging and cleaning the Slovene semantic lexicon slowNet, where we collaborated in the development of two tools: slowTool, a browser and editor for slowNet, and slowCrowd, a crowdsourcing tool for lexicon cleaning. In 2012 we started with work on implementing web services to linguistically annotated texts; so far we have implemented part-of-speech tagging and lemmatization for Slovene and English in the workflow construction environments Orange4WS and CloudFlows.

Launch of a novel, public sentiment monitoring system in collaboration with Gama System and POP TV, aired in prime time during the 2012 presidential elections.



Figure 3: Twitter sentiment about candidates at the 2012 Slovenian presidential elections shown on POP TV. Source: <http://www.predsedniskevotivte.si/>, Gama System in collaboration with Jožef Stefan Institute, Slovenia.

We collaborated in the work of the Slovene Institute of Standardization, as the Slovene delegates of ISO/TC37/SC4 "Terminology and Other Language and Content Resources / Language Resources Management" by taking part in the meetings of ISO TC 37 and in reviewing, translating and approving Slovene standards from this field. We were active in the preparation of the National Program for Language Policy 2012-2016 and in taking the steps necessary for Slovenia to join the research infrastructure CLARIN (Common Language Resources and Technology Infrastructure) at the Slovene Ministry of Culture.

In the area of decision support our long-term goal is to develop methods and techniques of decision modeling, support them with software and inte-

grate them with data-mining systems. In 2012 we developed a new version (3.04) of the computer program for multi-attribute decision-making DEXi, which brings better reporting capabilities and completely re-implemented JDEXi, a java library for the evaluation of decision alternatives. We also improved existing and developed new methods for the ranking of alternatives in qualitative multi-attribute models, based on copulas, which improve the sensitivity of decision models and alleviate some drawbacks of existing methods. In the framework of the 7FP EU project e-LICO, we developed a multi-attribute model for the evaluation of workflows in Rapid Miner, a well-known data-mining suite. In the 7FP project FIRST, we started the development of evaluation models for banking and finance. We developed a model for the assessment of roof coverings, which is methodologically interesting for its explicit modeling of context-based dependencies. Our results on six decision-support models related to growing and using genetically-modified crops, which were achieved in the already finished EU 7FP project Co-Extra, were published in 2012 in a Wiley-Blackwell book.

We participated in the project EVADIFF (Evaluation et de développement et modèles outils d'aide à la décision utilisés pour la Prévention des pollutions diffuses par les produits phytopharmaceutiques), commissioned by ARVALIS Institut du Végétal, France, where we develop a decision-support system for the selection of mitigation measures for the protection of surface waters from pollution by phytopharmaceuticals.

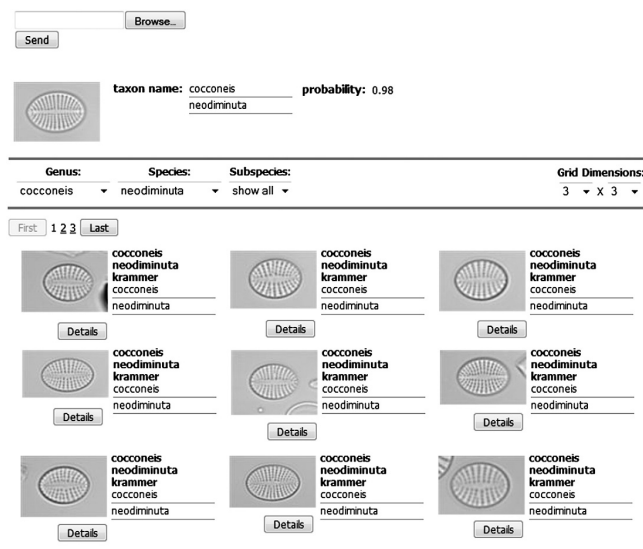


Figure 4: User interface for a system for the automatic annotation of microscopic diatom images. The query image and the reported identification results (taxon name and probability) are shown in the upper part. The bottom part shows the reference images for the identified taxon name that are present in the database.

Some outstanding publications in the past year

1. Stojanova, D., Ceci, M., Appice, A., Džeroski, S.: Network regression with predictive clustering trees. Data mining and knowledge discovery, 2012, vol. 25, no. 2, pp. 378–413
2. Grčar, M., Trdin, N., Lavrač, N.: A methodology for mining document-enriched heterogeneous information networks. Comput. j., [in press] 2012, 15 pages
3. Aho, T., Ženko, B., Džeroski, S., Elomaa, T.: Multi-target regression with rule ensembles. J. mach. learn. res., [Print ed.], 2012, vol. 13, pp. 2367–2407
4. Miljković, D., Stare, T., Mozetič, I., Podpečan, V., Petek, M., Witek, K., Dermastia, M., Lavrač, N., Gruden, K.: Signalling network construction for modelling plant defence response. PloS one, 2012, vol. 7, no. 12, pp. e51822–e51822-18. <http://ponta.ijs.si/mozetic/papers/Miljetal-signet-PlosONE-12.pdf>
5. Logar Berginc, N., Grčar, M., Brakus, M., Erjavec, T., Arhar Holdt, Š., Krek, S.: Korpusi slovenskega jezika Gigafida, KRES, ccGigafida in ccKRES: gradnja, vsebina, uporaba, (Zbirka Sporazumevanje). 1. izd. Ljubljana: Trojina, zavod za uporabno slovenistiko: Fakulteta za družbene vede, 2012. 208 pages, ilustr. ISBN 978-961-92983-6-7. ISBN 978-961-235-596-8
6. Mileva-Boshkoska, B., Bohanec, M.: A method for ranking non-linear qualitative decision preferences using copulas. International journal of decision support system technology, [in press] 2012, 17 pages

Nada Lavrač gave an invited lecture at The 25th IEEE International Symposium on Computer-Based Medical Systems - CBMS 2012, held in Rome in June 2012.

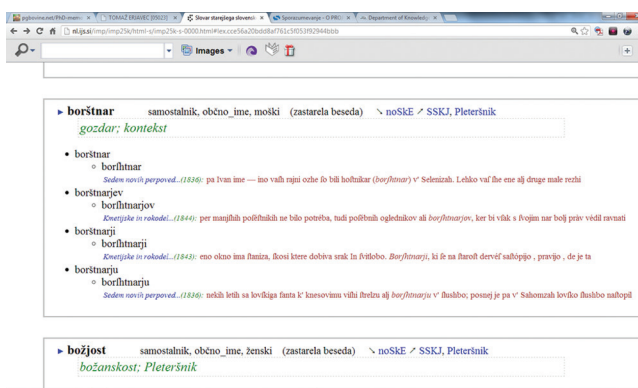


Figure 5: The first digital dictionary of historical Slovene, containing over 20,000 entries and available from <http://nl.ijs.si/imp/>, can be used by linguists, as well as for applications in human language technologies, e.g., to support full-text search in digital libraries of Slovenian cultural heritage.

Organization of conferences, congresses and meetings

1. The 22nd International Conference on Inductive Logic Programming - ILP 2012, Dubrovnik, Croatia, 17.–19. 9. 2012
2. Subconference: Intelligent Systems, and Conference on 100 Years of Alan Turing and 20 Years Of SLAIS, Information Society 2012, Ljubljana, Slovenia 8.–12. 10. 2012
3. Project meeting of European project REWIRE, Ljubljana, Slovenia, 17.–18. 9. 2012

INTERNATIONAL PROJECTS

1. EVADIFF: Evaluation of existing models and development of new decision-making tools to prevent diffuse pollution caused by plant protection products
Arvalis - Institut du Végétal
Prof. Marko Debeljak
2. 7. FP - PHAGOSYS: Systems biology of phagosome formation and maturation, modulation by intracellular pathogens
European Commission
Prof. Sašo Džeroski
3. 7. FP - ENVISION: Environmental services infrastructures with ontologies
European Commission
Miha Grčar, Prof. Nada Lavrač
4. 7. FP - IMPACT: Improving access to text
European Commission
Asst. Prof. Tomaž Erjavec
7. FP - FIRST: Large scale information extraction and integration infrastructure for supporting financial decision making
European Commission
Miha Grčar, Prof. Nada Lavrač
5. 7. FP - e-LICO: E-laboratory for collaborative interdisciplinary research in data mining and data intensive sciences
European Commission
Prof. Nada Lavrač, Asst. Prof. Martin Žnidaršič
6. 7. FP - SUMO: Supermodeling by combining imperfect models
European Commission
Prof. Sašo Džeroski
7. FP - FOC-II: Forecasting financial crises
European Commission
Dr. Igor Mozetič, Miha Grčar
7. FP - REWIRE: Rehabilitative way-out in responsive home environments
European Commission
Prof. Sašo Džeroski, Asst. Prof. Bernard Ženko
7. FP - MUSE: Machine understanding for interactive storytelling
European Commission
Prof. Nada Lavrač
9. COST IC1002, MUMIA: Multilingual and multifaceted interactive information access
COST Office
Dr. Igor Mozetič, Asst. Prof. Tomaž Erjavec

10. The European network on word structure
ESF - European Science Foundation
Asst. Prof. Tomaž Erjavec
11. Identifying optimal management strategies for biodiversity and related ecosystem services on private forests
Slovenian Research Agency
Prof. Marko Debeljak

RESEARCH PROGRAM

1. Knowledge Technologies
Prof. Nada Lavrač

R & D GRANTS AND CONTRACTS

1. Systemic biology approaches to analyzing interactions between pathogens and plants
Prof. Nada Lavrač
2. Slovene translation studies - resources and research
Asst. Prof. Tomaž Erjavec
3. Growth and defense trade-offs in multitrophic interaction between potato and its two major pests
Prof. Nada Lavrač
4. The leading humanists in the Slovenian territory between the 16th and mid-19th centuries and their social and cultural environment
Asst. Prof. Tomaž Erjavec
5. Data mining for integrative data analysis in systemic biology
Prof. Sašo Džeroski
6. Semantic rule discovery in the context of Web services
Prof. Nada Lavrač
7. Ecological restoration of natural disturbances in forests
Prof. Marko Debeljak
8. Google digital humanities award for developing language models for historical Slovene
Asst. Prof. Tomaž Erjavec
9. 22nd International conference on inductive logic programming Dubrovnik - ILP 2012, 17.-19.9.2012, Dubrovnik
Prof. Nada Lavrač

VISITORS FROM ABROAD

1. Dr. Ivica Dimitrovski, Faculty of Electrical Engineering and Information Technologies, University Ss. Cyril and Methodius, Skopje, Macedonia, 8.-29. 1. 2012
2. Dr. Gjorgji Madjarov, Faculty of Electrical Engineering and Information Technologies, University Ss. Cyril and Methodius, Skopje, Macedonia, 8.-29. 1. 2012
3. Dr. Nikola Ljubešić, Univerza v Zagrebu, Zagreb, Croatia, 6.-9. 2. 2012
4. Prof. dr. Suzana Loškoska, Faculty of Electrical Engineering and Information Technologies, University Ss. Cyril and Methodius, Skopje, Macedonia, 1. 3.-1. 10. 2012
5. Akad. prof. dr. Ljupčo Kocarev, Macedonian Academy of Sciences and Arts, Skopje, Macedonia, 11.-16. 3. 2012
6. Dr. Michelangelo Ceci, Università degli Studi di Bari, Bari, Italy, 23.-25. 4. 2012
7. Dr. Michelangelo Puliga, ETH, Zurich, Switzerland, 15.-18. 5. 2012, 18. 12. 2012
8. Dr. Dragan Gamberger, Rudjer Bošković Institute, Zagreb, Croatia, 15.-18. 5. 2012
9. Dr. Tomislav Šmuc, Rudjer Bošković Institute, Zagreb, Croatia, 15.-18. 5. 2012, 15. 11. 2012
10. Nino Antulov Fantulin, Rudjer Bošković Institute, Zagreb, Croatia, 15.-18. 5. 2012, 18. 12. 2012
11. Matija Piškorec, Rudjer Bošković Institute, Zagreb, Croatia, 15.-18. 5. 2012, 15. 11. 2012
12. Marko Popović, Rudjer Bošković Institute, Zagreb, Croatia, 15.-18. 5. 2012, 4.-5. 11. 2012, 4.-5. 12. 2012
13. Dražen Lučanin, Rudjer Bošković Institute, Zagreb, Croatia, 15.-18. 5. 2012
14. Prof. dr. Benedict Brors, German Cancer Research Center - DKFZ, Heidelberg, Germany, 25.-27. 7. 2012
15. Dr. Florence Leprince, ARVALIS - Institut du végétal, Montardon, France, 6.-8. 9. 2012, 7.-8. 11. 2012
16. Dr. Barry Hardy, Douglas Connect, Zeiningen, Switzerland, 5. 5. 2012
17. Richard Wheeler, University of Edinburgh, Edinburgh, Scotland, 4.-9. 9. 2012
18. Prof. dr. Hiroshi Motoda, AFOSR/AOARD in University of Osaka, Osaka, Japan, 19.-23. 9. 2012
19. Prof. dr. João Gama, University of Porto, Laboratory of Artificial Intelligence and Decision Support, and Faculty of Economics, Porto, Portugal, 8.-14. 10. 2012
20. Dr. Nataša Pržulj, Imperial College, London, UK, 6. 11. 2012
21. Dr. Benoit Real, ARVALIS Institut du végétal, Paris, France, 7.-8. 11. 2012

STAFF

Researchers

1. Prof. Marko Bohanec
2. Prof. Bojan Cestnik*
3. Prof. Marko Debeljak
4. Prof. Sašo Džeroski
5. Asst. Prof. Tomaž Erjavec
6. **Prof. Nada Lavrač, Head**
7. Prof. Tanja Urbančič*

Postdoctoral associates

8. Dr. Dragi Kocev
9. Dr. Petra Kralj Novak
10. Dr. Panče Panov
11. Dr. Ivica Slavkov

12. Dr. Aneta Trajanov
13. Asst. Prof. Bernard Ženko
14. Asst. Prof. Martin Žnidaršič

Postgraduates

15. Darko Čerepnalkoski, B. Sc.
16. Miha Grčar, B. Sc.
17. Dr. Elena Ikonomovska
18. Matjaž Juršič, B. Sc.
19. Janez Kranjc, B. Sc.
20. Biljana Mileva Boshkoska, M. Sc.
21. Matic Perovšek, B. Sc.
22. Vid Podpečan, B. Sc.
23. Senja Pollak, B. Sc.

24. Nikola Simidjievski, B. Sc.

25. Borut Sluban, B. Sc.

26. Nejc Trdin, B. Sc.

27. Anže Vavpetič, B. Sc.

Technical officers

28. Marko Brakus, B. Sc., left 01.05.12

29. Dr. Igor Mozetič

Technical and administrative staff

30. Tina Anžič, B. Sc.

31. Milica Bauer, B. Sc.

32. Dr. France Dacar, retired 29.07.12

Note:

* part-time JSI member

BIBLIOGRAPHY

ORIGINAL SCIENTIFIC ARTICLE

1. Timo Aho, Bernard Ženko, Sašo Džeroski, Tapio Elomaa, "Multi-target regression with rule ensembles", *J. mach. learn. res.*, vol. 13, pp. 2367-2407, 2012.
2. Darko Čerepnalkoski, Katerina Taškova, Ljupčo Todorovski, Nataša Atanasova, Sašo Džeroski, "The influence of parameter fitting methods on model structure selection in automated modeling of aquatic ecosystems", In: Proceedings of the 7th ECEM, European Conference on Ecological Modelling, 30 May - 2 June 2011, Riva el Garda, Italy, *Ecol. Model.*, vol. 245, pp. 136-166, 2012.
3. Marko Debeljak, Aneta Trajanov, Daniela Stojanova, Florence Leprince, Sašo Džeroski, "Using relational decision trees to model out-crossing rates in a multi-field setting", In: Proceedings of the 7th ECEM, European Conference on Ecological Modelling, 30 May - 2 June 2011, Riva el Garda, Italy, *Ecol. Model.*, vol. 245, pp. 75-83, 2012.
4. Ivica Dimitrovski, Dragi Kocev, Suzana Loskovska, Sašo Džeroski, "Hierarchical classification of diatom images using ensembles of predictive clustering trees", *Ecological informatics*, vol. 7, no. 1, pp. 19-29, 2012.
5. Tomaž Erjavec, "Jezikoslovni viri starejše slovenščine", In: *Ljubljana v BiTiH - BiTi v Ljubljani: prispevki iz prvega ljubljanskega kongresa digitalizacije kulturne dediščine = papers from the first Slovenian congress for digitisation of cultural heritage: tematska številka*, (Knjižnica, 56, 3), Ines Vodopivec, ed., Ljubljana, Zveza bibliotekarskih društev Slovenije, Narodna in univerzitetna knjižnica, 2012, pp. 205-221.
6. Tomaž Erjavec, "MULTEXT-East: morphosyntactic resources for Central and Eastern European languages", *Language resources and evaluation*, vol. 46, no. 1, pp. 131-142, 2012.
7. Frieder Graef et al. (38 authors), "A framework for a European network for a systematic environmental impact assessment of genetically modified organisms (GMO)", *BioRisk (Print)*, vol. 7, pp. 73-97, 2012.
8. Miha Grčar, Vid Podpečan, Borut Sluban, Igor Mozetič, "Ontology querying support in semantic annotation process", In: PRICAI 2012: trends in artificial intelligence: 12th Pacific Rim International Conference, Kuching, Malaysia, September 3-7, 2012: proceedings, *Lecture notes in computer science*, vol. 7458, pp. 76-87, 2012.
9. Janez Kranjc, Vid Podpečan, Nada Lavrač, "CloudFlows: a cloud based scientific workflow platform", In: Machine learning and knowledge discovery in databases: European conference, ECML PKDD 2012 Bristol, UK, September 24-28, 2012: part II: proceedings, *Lecture notes in computer science*, vol. 7523, pp. 816-819, 2012.
10. Gjorgji Madžarov, Dejan Gjorgjevič, Sašo Džeroski, "Two stage architecture for multi-label learning", *Pattern recogn.*, vol. 45, no. 3, pp. 1019-1034, 2011.
11. Gjorgji Madžarov, Dragi Kocev, Dejan Gjorgjevič, Sašo Džeroski, "An extensive experimental comparison of methods for multi-label learning", In: Proceedings of the Iberian Conference on Pattern Recognition and Image Analysis (IbPRIA'2011), 8-10 June 2011, Las Palmas de Gran Canaria, Spain, *Pattern recognition*, vol. 45, no. 9, pp. 3084-3104, 2012.
12. Dragana Miljković, Tjaša Stare, Igor Mozetič, Vid Podpečan, Marko Petek, Kamil Witek, Marina Dermastia, Nada Lavrač, Kristina Gruden, "Signalling network construction for modelling plant defence response", *PLoS one*, vol. 7, no. 12, pp. e51822-1e51822-18, 2012.
13. Patrik Mouron et al. (22 authors), "Sustainability assessment of crop protection systems: sustainOS methodology and its application for apple orchards", *Agric. syst.*, vol. 113, pp. 1-15, 2012.
14. Ingrid Petrič, Bojan Cestnik, Nada Lavrač, Tanja Urbančič, "Outlier detection in cross-context link discovery for creative literature mining", *Comput. j.*, vol. 55, no. 1, pp. 47-61, 2012.
15. Vid Podpečan, Monika Žáková, Nada Lavrač, "Orange4WS environment for service-oriented data mining", *Comput. j.*, vol. 55, no. 1, pp. 82-98, 2012.
16. Senja Pollak, Nejc Trdin, Anže Vavpetič, Tomaž Erjavec, "NLP web services for Slovene and English: morphosyntactic tagging, lemmatisation and definition extraction", *Informatica (Ljublj.)*, vol. 36, no. 4, pp. 441-449, 2012.
17. Daniela Stojanova, Michelangelo Ceci, Annalisa Appice, Sašo Džeroski, "Network regression with predictive clustering trees", In: *Proceedings of the ECML PKDD 2011, European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases*, 5-9 September 2011, Athens, Greece, *Data mining and knowledge discovery* vol. 25, no. 2, pp. 378-413, 2012.
18. Daniela Stojanova, Andrej Kobler, Peter Ogrinc, Bernard Ženko, Sašo Džeroski, "Estimating the risk of fire outbreaks in the natural environment", In: Proceedings of the ECML PKDD 2011, European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases, 5-9 September 2011, Athens, Greece, *Data mining and knowledge discovery* vol. 24, no. 2, pp. 411-442, 2012.
19. Katerina Taškova, Jurij Šilc, Nataša Atanasova, Sašo Džeroski, "Parameter estimation in a nonlinear dynamic model of an aquatic ecosystem with meta-heuristic optimization", *Ecol. model.*, vol. 226, no. 1, pp. 36-61, 2012.
20. Anže Vavpetič, Vid Podpečan, Stijn Meganck, Nada Lavrač, "Explaining subgroups through ontologies", In: PRICAI 2012: trends in artificial intelligence: 12th Pacific Rim International Conference, Kuching, Malaysia, September 3-7, 2012: proceedings, *Lecture notes in computer science*, vol. 7458, pp. 625-636, 2012.

REVIEW ARTICLE

1. Matija Ogrin, Jan Jona Javoršek, Tomaž Erjavec, "Register slovenskih rokopisov 17. in 18. stoletja: repozitorij, digitalna knjižnica in raziskovalno okolje: repository, digital library and research environment", In: *Ljubljana v BiTiH - BiTi v Ljubljani: prispevki iz prvega ljubljanskega kongresa digitalizacije kulturne dediščine = papers from the first Slovenian congress for digitisation of cultural heritage: tematska številka*, (Knjižnica, 56, 3), Ines Vodopivec, ur., Ljubljana, Zveza bibliotekarskih društev Slovenije, Narodna in univerzitetna knjižnica, 2012, pp. 161-173.

SHORT SCIENTIFIC ARTICLE

1. Dragi Kocev, "Ensembles for predicting structured outputs", *Informatica (Ljublj.)*, vol. 36, no. 1, pp. 113-114, 2012.

PUBLISHED SCIENTIFIC CONFERENCE CONTRIBUTION (INVITED LECTURE)

1. Nada Lavrač, "Advances in data mining for biomedical research", In: *CBMS 2012, The 25th IEEE International Symposium on Computer-Based Medical System*, CBMS 2012, June 20-22, Rome, Italy, [Piscataway], Institute of Electrical and Electronics Engineers, = IEEE, cop. 2012, 5 pp.

PUBLISHED SCIENTIFIC CONFERENCE CONTRIBUTION

1. Marko Bohanec, Vladislav Rajkovič, Ivan Bratko, Blaž Zupan, Martin Žnidaršič, "DEX methodology: thirty three years of qualitative multi-attribute modeling", In: *Zbornik 15. mednarodne multikonference Informacijska družba - IS 2012, 8.-12. oktober 2012, Ljubljana, Slovenia:*

- zvezek A: volume A, (Informacijska družba), Marko Bohanec, ed., Matjaž Gams, ed., Dunja Mladenec, ed., Marko Grobelnik, ed., Marjan Heričko, ed., Urban Kordeš, ed., Maja Smrdu, ed., Olga Markič, ed., Zvezdan Pirtošek, ed., Jadran Lenarčič, ed., Leon Žlajpah, ed., Andrej Gams, ed., Vladislav Rajkovič, ed., Tanja Urbančič, ed., Mojca Bernik, ed., Ljubljana, Institut Jožef Stefan, 2012, pp. 31-34.
2. Bojan Cestnik, Alenka Kern, "Social computing potential for citizen engagement in public sector services", In: *CeDEM12: proceedings of the International Conference for E-democracy and Open Government, 3-4 May 2012, Krems, Austria*, Peter Parycek, ed., Noella Edelmann, ed., Krems, Donau-Universität, 2012, pp. 303-307.
 3. Sašo Džeroski, "Machine learning for systems biosciences", In: *Zbornik 15. mednarodne multikonference Informacijska družba - IS 2012, 8.-12. oktober 2012, Ljubljana, Slovenia: zvezek A: volume A*, (Informacijska družba), Marko Bohanec, ed., Matjaž Gams, ed., Dunja Mladenec, ed., Marko Grobelnik, ed., Marjan Heričko, ed., Urban Kordeš, ed., Maja Smrdu, ed., Olga Markič, ed., Zvezdan Pirtošek, ed., Jadran Lenarčič, ed., Leon Žlajpah, ed., Andrej Gams, ed., Vladislav Rajkovič, ed., Tanja Urbančič, ed., Mojca Bernik, ed., Ljubljana, Institut Jožef Stefan, 2012, zv. A, pp. 37-40.
 4. Tomaž Erjavec, "The goo300k corpus of historical Slovene", In: *LREC 2012: proceedings, 8th International Conference on Language Resources and Evaluation, 21-27 May 2012, Istanbul, Turkey*, Istanbul, ELRA, 2012, pp. 2257-2260.
 5. Tomaž Erjavec, "Jezikovni viri starejše slovenščine IMP: zbirka besedil, korpus, slovar", In: *Zbornik Osmе konference Jezikovne tehnologije, 8. do 12. oktober 2012, [Ljubljana, Slovenia]: zbornik 15. mednarodne multikonference Informacijska družba - IS 2012, zvezek C: proceedings of the 15th International Multiconference Information Society - IS 2012, volume C*, (Informacijska družba), Tomaž Erjavec, ed., Jerneja Žganec Gros, ed., Ljubljana, Institut Jožef Stefan, 2012, pp. 52-56.
 6. Tomaž Erjavec, Nataša Logar Berginc, "Referenčni korpusi slovenskega jezika (cc)Gigafida in (cc)KRES", In: *Zbornik Osmе konference Jezikovne tehnologije, 8. do 12. oktober 2012, [Ljubljana, Slovenia]: zbornik 15. mednarodne multikonference Informacijska družba - IS 2012, zvezek C: proceedings of the 15th International Multiconference Information Society - IS 2012, volume C*, (Informacijska družba), Tomaž Erjavec, ed., Jerneja Žganec Gros, ed., Ljubljana, Institut Jožef Stefan, 2012, pp. 57-62.
 7. Darja Fišer, Jernej Novak, Tomaž Erjavec, "sloWNet 3.0: development, extension and cleaning", In: *6th International Global Wordnet Conference: [proceedings]*, [S. l.], The Global WordNet Association, [2012], pp. 113-117.
 8. Miha Grčar, Simon Krek, Kaja Dobrovoljc, "Obeliks: statistični oblikoskladenski označevalnik in lematizator za slovenski jezik: statistical morphosyntactic tagger and lemmatizer for Slovene", In: *Zbornik Osmе konference Jezikovne tehnologije, 8. do 12. oktober 2012, [Ljubljana, Slovenia]: zbornik 15. mednarodne multikonference Informacijska družba - IS 2012, zvezek C: proceedings of the 15th International Multiconference Information Society - IS 2012, volume C*, (Informacijska družba), Tomaž Erjavec, ed., Jerneja Žganec Gros, ed., Ljubljana, Institut Jožef Stefan, 2012, pp. 89-94.
 9. Matjaž Juršič, Bojan Cestnik, Tanja Urbančič, Nada Lavrač, "Cross-domain literature mining: Finding bridging concepts with CrossBee", In: *Proceedings of the Third International Conference on Computational Creativity, ICC3 2012, May 30 - June 1, 2012, Dublin, Ireland*, Mary Lou Maher, ed., Dublin, University College, 2012, pp. 33-40.
 10. Tom Kentner, Tomaž Erjavec, Maja Žorga Dulmin, Darja Fišer, "Lexicon construction and corpus annotation of historical language with CoBaLT editor", In: *Proceedings of the 6th Workshop on Language Technology for Cultural Heritage, Social Sciences, and Humanities (LaTeCH 2012)*, 13th Conference on the European Chapter of the Association for Computational Linguistics, Avignon, April 23-17 2012, Avignon, ACL, 2012.
 11. Janez Kranjc, Vid Podpečan, Nada Lavrač, "Knowledge discovery using a service oriented web application", In: *DigitalWorld 2012: January 30 - February 4, 2012 - Valencia, Spain*, [S. l.], International Academy, Research, and Industry Association, = IARIA, cop. 2012, pp. 82-86.
 12. Vladimir Kuzmanovski, Sašo Džeroski, Marko Debeljak, "Integration of structured expert knowledge", In: *Zbornik, 4. študentska konferenca Mednarodne podiplomske šole Jožefa Stefana = 4th Jožef Stefan International Postgraduate School Students Conference*, 25. maj 2012, Ljubljana, Slovenija, Dejan Petelin, ed., Aleš Tavčar, ed., Boštjan Kaluža, ed., Ljubljana, Mednarodna podiplomska šola Jožefa Stefana, 2012, pp. 137-143.
 13. Nada Lavrač, Petra Kralj Novak, "Relational and semantic data mining for biomedical research", In: *Zbornik 15. mednarodne multikonference Informacijska družba - IS 2012, 8.-12. oktober 2012, Ljubljana, Slovenia: zvezek A: volume A*, (Informacijska družba), Marko Bohanec, ed., Matjaž Gams, ed., Dunja Mladenec, ed., Marko Grobelnik, ed., Marjan Heričko, ed., Urban Kordeš, ed., Maja Smrdu, ed., Olga Markič, ed., Zvezdan Pirtošek, ed., Jadran Lenarčič, ed., Leon Žlajpah, ed., Andrej Gams, ed., Vladislav Rajkovič, ed., Tanja Urbančič, ed., Mojca Bernik, ed., Ljubljana, Institut Jožef Stefan, 2012, zv. A, pp. 27-30.
 14. Sandi Marinič, Marko Bohanec, "Večparametrsko vrednotenje variant v odvisnosti od konteksta: model za vrednotenje strešnih kritin", In: *Zbornik 15. mednarodne multikonference Informacijska družba - IS 2012, 8.-12. oktober 2012, Ljubljana, Slovenia: zvezek A: volume A*, (Informacijska družba), Marko Bohanec, ed., Matjaž Gams, ed., Dunja Mladenec, ed., Marko Grobelnik, ed., Marjan Heričko, ed., Urban Kordeš, ed., Maja Smrdu, ed., Olga Markič, ed., Zvezdan Pirtošek, ed., Jadran Lenarčič, ed., Leon Žlajpah, ed., Andrej Gams, ed., Vladislav Rajkovič, ed., Tanja Urbančič, ed., Mojca Bernik, ed., Ljubljana, Institut Jožef Stefan, 2012, zv. A, pp. 76-79.
 15. Biljana Mileva-Boshkoska, Marko Bohanec, "Ranking of qualitative decision options using copulas", In: *Operations research proceedings 2011: selected papers of the International Conference on Operations Research (OR 2011), August 30 - September 2, 2011, Zurich, Switzerland*, (Operations research proceedings), Diethard Klatte, ed., Berlin, Heidelberg, Springer, cop. 2012, pp. 103-108.
 16. Biljana Mileva-Boshkoska, Marko Bohanec, Martin Žnidaršič, "Experimental evaluation of methods for ranking qualitatively assessed data-mining workflows", In: *Fusing decision support systems into the fabric of the context: [presented at 16th IFIP WG8.3 International Conference on Decision Support Systems, June 28-30 2012, Anáissos, Greece]*, (Frontiers in artificial intelligence and applications, vol. 238), Ana Respício, ed., Frada Burstein, ed., Amsterdam, IOS Press, cop. 2012, pp. 175-184.
 17. Dragana Miljković, Matjaž Depolli, Igor Mozetič, Nada Lavrač, Tjaša Stare, Marko Petek, Kristina Gruden, "Constraint-driven optimization of plant defense model parameters", In: *Proceedings, 2012 IEEE International Conference on Bioinformatics and Biomedicine Workshops (BIBMW)*, 4-7 October 2012, Philadelphia, Jean Gao, ed., Danvers, Institute of Electrical and Electronics Engineers, 2012, pp. 570-574.
 18. Matic Perovšek, Nada Lavrač, Bojan Cestnik, "Visual divisive hierarchical clustering using k-means", In: *Zbornik 15. mednarodne multikonference Informacijska družba - IS 2012, 8.-12. oktober 2012, Ljubljana, Slovenia: zvezek A: volume A*, (Informacijska družba), Marko Bohanec, ed., Matjaž Gams, ed., Dunja Mladenec, ed., Marko Grobelnik, ed., Marjan Heričko, ed., Urban Kordeš, ed., Maja Smrdu, ed., Olga Markič, ed., Zvezdan Pirtošek, ed., Jadran Lenarčič, ed., Leon Žlajpah, ed., Andrej Gams, ed., Vladislav Rajkovič, ed., Tanja Urbančič, ed., Mojca Bernik, ed., Ljubljana, Institut Jožef Stefan, 2012, zv. A, pp. 193-196.
 19. Senja Pollak, Nejc Trdin, Anže Vavpetič, Tomaž Erjavec, "A web service implementation of linguistic annotation for Slovene and English", In: *Zbornik Osmе konference Jezikovne tehnologije, 8. do 12. oktober 2012, [Ljubljana, Slovenia]: zbornik 15. mednarodne multikonference Informacijska družba - IS 2012, zvezek C: proceedings of the 15th International Multiconference Information Society - IS 2012, volume C*, (Informacijska družba), Tomaž Erjavec, ed., Jerneja Žganec Gros, ed., Ljubljana, Institut Jožef Stefan, 2012, pp. 157-162.
 20. Senja Pollak, Anže Vavpetič, Janez Kranjc, Nada Lavrač, Špela Vintar, "NLP workflow for on-line definition extraction from English and Slovene text corpora", In: *Empirical methods: proceedings of the Conference on Natural Language Processing 2012*, (Scientific series of the ŌGAI, volume 5), 11th Conference on Natural Language Processing (KONVENS) [September 19-21, 2012, Vienna, Austria], Jeremy Jancsary, ed., Wien, ŌGAI, = Österreichischen Gesellschaft für Artificial Intelligence, 2012, pp. 53-60.
 21. Borut Sluban, Senja Pollak, Roel Coesemans, Nada Lavrač, "Irregularity detection in categorized document corpora", In: *LREC 2012: proceedings*, 8th International Conference on Language Resources and Evaluation, 21-27 May 2012, Istanbul, Turkey, Istanbul, ELRA, 2012, pp. 1598-1603.
 22. Jasmina Smailović, Miha Grčar, Martin Žnidaršič, "Sentiment analysis on tweets in a financial domain", In: *Zbornik, 4. študentska konferenca Mednarodne podiplomske šole Jožefa Stefana = 4th Jožef Stefan International Postgraduate School Students Conference*, 25. maj 2012, Ljubljana, Slovenija, Dejan Petelin, ed., Aleš Tavčar, ed., Boštjan Kaluža, ed., Ljubljana, Mednarodna podiplomska šola Jožefa Stefana, 2012, pp. 169-175.
 23. Jasmina Smailović, Senja Pollak, "Topic ontology construction from English and Slovene language technologies corpora", In: *Zbornik Osmе konference Jezikovne tehnologije, 8. do 12. oktober 2012, [Ljubljana,*

- Slovenia]: zbornik 15. mednarodne multikonference Informacijska družba - IS 2012, zvezek C: proceedings of the 15th International Multiconference Information Society - IS 2012, volume C, (Informacijska družba), Tomaž Erjavec, ed., Jerneja Žganec Gros, ed., Ljubljana, Institut Jožef Stefan, 2012, pp. 173-178.
24. Larisa N. Soldatova, Sašo Džeroski, Panče Panov, "Relation for information entities ... [et al.]", In: *The 15th Annual Bio-Ontologies Meeting, July 13-14, 2012, Long Beach, Ca, USA*, The 15th Annual Bio-Ontologies Meeting, July 13-14, 2012, Long Beach, Ca, USA, [S. l., s. n.], 2012, pp. 1-4.
 25. Daniela Stojanova, Michelangelo Ceci, Annalisa Appice, Sašo Džeroski, "Network regression with predictive clustering trees", In: *Proceedings of the ECML PKDD 2011, European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases, 5-9 September 2011, Athens, Greece*, (Data mining and knowledge discovery, Vol. 25, no. 2, 2012), Dimitrios Gunopulos, ed., Donato Malerba, ed., Michalis Vazirgiannis, ed., Boston, Dordrecht, London, Kluwer, 2012, vol. 25, no. 2, pp. 378-413, 2012.
 26. Tadej Štajner, Tomaž Erjavec, Simon Krek, "Razpoznavanje imenskih entitet v slovenskem jeziku", In: *Zbornik Osmе konference Jezikovne tehnologije, 8. do 12. oktobra 2012, [Ljubljana, Slovenija]: zbornik 15. mednarodne multikonference Informacijska družba - IS 2012, zvezek C: proceedings of the 15th International Multiconference Information Society - IS 2012, volume C*, (Informacijska družba), Tomaž Erjavec, ed., Jerneja Žganec Gros, ed., Ljubljana, Institut Jožef Stefan, 2012, pp. 191-196.
 27. Jovan Tanevski, Nikola Simidjievski, Sašo Džeroski, "Biocircuit design with equation discovery", In: *Pre workshop proceedings, LDSB'12, ECML-PKDD 2012, Workshop on Learning and Discovery in Symbolic Systems Biology*, in collaboration with the European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases, September 24, 2012 Bristol, UK, [S. l., s. n.], 2012, pp. 2-16.
 28. Aleš Tavčar, Darja Fišer, Tomaž Erjavec, "sloWCrowd: orodje za popraviljanje wordneta z izkoriščanjem moči množic", In: *Zbornik Osmе konference Jezikovne tehnologije, 8. do 12. oktobra 2012, [Ljubljana, Slovenija]: zbornik 15. mednarodne multikonference Informacijska družba - IS 2012, zvezek C: proceedings of the 15th International Multiconference Information Society - IS 2012, volume C*, (Informacijska družba), Tomaž Erjavec, ed., Jerneja Žganec Gros, ed., Ljubljana, Institut Jožef Stefan, 2012, pp. 197-202.
 29. Nejc Trdin, Marko Bohanec, "Extending the multi-criteria decision making method DEX", In: *Zbornik, 4. študentska konferenca Mednarodne podiplomske šole Jožefa Stefana = 4th Jožef Stefan International Postgraduate School Students Conference*, 25. maj 2012, Ljubljana, Slovenija, Dejan Petelin, ed., Aleš Tavčar, ed., Boštjan Kaluža, ed., Ljubljana, Mednarodna podiplomska šola Jožefa Stefana, 2012, pp. 182-187.
 30. Ines Vodopivec, Tomaž Erjavec, Daša Pokorn, Alenka Kavčič-Čolič, "Optical character recognition of historical texts: end-user focused research for Slovenian books and newspapers from the 18th and 19th Century", *Pregl. Nac. cent. digit.*, [Br.] 21, pp. 117-126, 2012.
 31. Martin Žnidaršič, Marko Bohanec, Nejc Trdin, "Qualitative assessment of data-mining workflows", In: *Fusing decision support systems into the fabric of the context: [presented at 16th IFIP WG8.3 International Conference on Decision Support Systems, June 28-30 2012, Anáissos, Greece]*, (Frontiers in artificial intelligence and applications, vol. 238), Ana Respício, ed., Frada Burstein, ed., Amsterdam, IOS Press, cop. 2012, pp. 75-88.
- an introduction to concept, algorithms, tools, and applications*, (Lecture notes in computer science, Vol. 7250), Michael R. Berthold, ed., Heidelberg [etc.], Springer, 2012, pp. 66-90.
4. Lado Kutnar, Andrej Kobler, Sašo Džeroski, "Napovedi spreminjanja deleža bukovih gozdov in obilja buke v spremenjenih okoljskih razmerah", In: *Bukovi gozdovi v Sloveniji: ekologija in gospodarjenje*, Andrej Bončina, ed., Ljubljana, Oddelek za gozdarstvo in obnovljive gozdne vire, Biotehniška fakulteta, 2012, pp. 259-270.
 5. Laura Langohr, Vid Podpečan, Marko Petek, Igor Mozetič, Kristina Gruden, "Contrast mining from interesting subgroups", In: *Bisociative knowledge discovery: an introduction to concept, algorithms, tools, and applications*, (Lecture notes in computer science, Vol. 7250), Michael R. Berthold, ed., Heidelberg [etc.], Springer, 2012, pp. 390-406.
 6. Dragana Miljković, Vid Podpečan, Miha Grčar, Kristina Gruden, Tjaša Stare, Marko Petek, Igor Mozetič, Nada Lavrač, "Modelling a biological system: network creation by triplet extraction from biological literature", In: *Bisociative knowledge discovery: an introduction to concept, algorithms, tools, and applications*, (Lecture notes in computer science, Vol. 7250), Michael R. Berthold, ed., Heidelberg [etc.], Springer, 2012, pp. 427-437.
 7. Igor Mozetič, Nada Lavrač, "Applications and evaluation: overview", In: *Bisociative knowledge discovery: an introduction to concept, algorithms, tools, and applications*, (Lecture notes in computer science, Vol. 7250), Michael R. Berthold, ed., Heidelberg [etc.], Springer, 2012, pp. 359-363.
 8. Igor Mozetič, Nada Lavrač, Vid Podpečan, Petra Kralj Novak, Helena Motaln, Marko Petek, Kristina Gruden, Hannu Toivonen, Kimmo Kulovesi, "Semantic subgroup discovery and cross-context linking for microarray data analysis", In: *Bisociative knowledge discovery: an introduction to concept, algorithms, tools, and applications*, (Lecture notes in computer science, Vol. 7250), Michael R. Berthold, ed., Heidelberg [etc.], Springer, 2012, pp. 379-389.
 9. Ingrid Petrič, Bojan Cestnik, Nada Lavrač, Tanja Urbančič, "Bisociative knowledge discovery by literature outlier detection", In: *Bisociative knowledge discovery: an introduction to concept, algorithms, tools, and applications*, (Lecture notes in computer science, Vol. 7250), Michael R. Berthold, ed., Heidelberg [etc.], Springer, 2012, pp. 313-324.
 10. Olivier Schmidt, Janez Kranjc, Igor Mozetič, Paul Thompson, Werner Dubitzky, "Bisociative exploration of biological and financial literature using clustering", In: *Bisociative knowledge discovery: an introduction to concept, algorithms, tools, and applications*, (Lecture notes in computer science, Vol. 7250), Michael R. Berthold, ed., Heidelberg [etc.], Springer, 2012, pp. 438-451.
 11. Borut Sluban, Matjaž Juršič, Bojan Cestnik, Nada Lavrač, "Exploring the power of outliers for cross-domain literature mining", In: *Bisociative knowledge discovery: an introduction to concept, algorithms, tools, and applications*, (Lecture notes in computer science, Vol. 7250), Michael R. Berthold, ed., Heidelberg [etc.], Springer, 2012, pp. 325-337.
 12. Daniela Stojanova, Marko Debeljak, Michelangelo Ceci, Annalisa Appice, Donato Malerba, Sašo Džeroski, "Dealing with spatial autocorrelation in gene flow modeling", In: *Models of the ecological hierarchy: from molecules to the ecosphere*, Ferenc Jordán, ed., Sven Erik Jørgensen, ed., [S. l.], Elsevier, 2012, pp. 35-49.

SCIENTIFIC MONOGRAPH

1. Johannes Fürnkranz, Dragan Gamberger, Nada Lavrač, *Foundations of rule learning*, (Cognitive technologies), Heidelberg [etc.], Springer, 2012.
2. Nataša Logar Berginc et al. (6 authors), *Korpusi slovenskega jezika Gigafida, KRES, ccGigafida in ccKRES: gradnja, vsebina, uporaba*, (Zbirka Sporazumevanje), 1. izd., Ljubljana, Trojina, zavod za uporabno slovenistiko, Fakulteta za družbene vede, 2012.

MENTORING

1. Elena Ikononovska, *Algorithms for learning regression trees and ensembles on evolving data streams*: doctoral dissertation, Ljubljana, 2012 (mentor Sašo Džeroski; co-mentor Joao Gama).
2. Panče Panov, *A modular ontology of data mining*: doctoral dissertation, Ljubljana, 2012 (mentor Sašo Džeroski).
3. Aleksander Pečkov, *A machine learning approach to polynomial regression*: doctoral dissertation, Ljubljana, 2012 (mentor Sašo Džeroski; co-mentor Ljupčo Todorovski).
4. Ivica Slavkov, *An evaluation method for feature rankings*: doctoral dissertation, Ljubljana, 2012 (mentor Sašo Džeroski).

INDEPENDENT SCIENTIFIC COMPONENT PART OR A CHAPTER IN A MONOGRAPH

1. Asta Gregorič, Boris Zmazek, Sašo Džeroski, Drago Torkar, Janja Vaupotič, "Radon as an earthquake precursor - methods for detecting anomalies", In: *Earthquake research and analysis: statistical studies, observations and planning*, Sebastiano D'Amico, ed., Rijeka, InTech, cop. 2011, pp. 179-196.
2. Matjaž Juršič, Bojan Cestnik, Tanja Urbančič, Nada Lavrač, "Bisociative literature mining by ensemble heuristics", In: *Bisociative knowledge discovery: an introduction to concept, algorithms, tools, and applications*, (Lecture notes in computer science, Vol. 7250), Michael R. Berthold, ed., Heidelberg [etc.], Springer, 2012, pp. 338-358.
3. Matjaž Juršič, Borut Sluban, Bojan Cestnik, Miha Grčar, Nada Lavrač, "Bridging concept identification for constructing information networks from text documents", In: *Bisociative knowledge discovery:*

5. Daniela Stojanova, *Considering autocorrelation in predictive models*: doctoral dissertation, Ljubljana, 2012 (mentor Sašo Džeroski).
6. Katerina Tashkova, *Parameter identification in nonlinear dynamic systems with meta-heuristic approaches*: doctoral dissertation, Ljubljana, 2012 (mentor Sašo Džeroski; co-mentor Jurij Šilc).
7. Jernej Vičič, *A fast implementation of rules based machine translation systems for similar natural languages*: doctoral dissertation, Ljubljana, 2012 (mentor Igor Kononenko; co-mentor Tomaž Erjavec).
8. Vladimir Kuzmanovski, *Integration of expert knowledge and predictive learning : modelling water flows in agriculture*: master's thesis, Ljubljana, 2012 (mentor Marko Debeljak; co-mentor Sašo Džeroski).
9. Marko Lazar, *Using web technologies for marketing communication in sport*: master's thesis, Nova Gorica, 2012 (mentors Bojan Cestnik, Tanja Urbančič).
10. Sandi Marinič, *A multi-attribute model for assessing roof coverings*: master's thesis, Nova Gorica, 2012 (mentor Marko Bohanec).