

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. The established areas of knowledge technologies include intelligent data analysis (machine learning, data mining, knowledge discovery in databases), text and web mining, the semantic web, social network analysis, language technologies and computational linguistics, as well as decision support and knowledge management. The research areas of the department also include Web 2.0, the management of virtual organizations, new media and e-science. In addition to knowledge technologies research, we also develop applications in environmental sciences and ecology, medicine and health care, biomedicine and genetics, as well as the economy and marketing.



Head:
Prof. Nada Lavrač

In 2008 the Knowledge Technologies research programme was evaluated as the best research programme in ICT (for the period 2004–2008) by the Slovenian Research Agency (ARRS). In terms of our collaboration in EU projects, our programme group is the most successful in Slovenia. In 2004–2008 we collaborated in 28 FP6 and FP7 projects, the most important being our successful coordination of the STREP project “Inductive Queries for Mining Patterns and Models” (IQ) and our collaboration in six large Integrated Projects (IPs) and two Networks of Excellence (NoEs).

In the area of **intelligent data analysis** we have developed a conceptual framework of supervised descriptive rule induction and have shown that methods for subgroup discovery, contrast set mining and emerging pattern mining represent special cases of supervised descriptive rule induction methods. An algorithm for closed-sets mining was published in the prestigious Journal of Machine Learning Research. A lemmatizer for Slovene, constructed by the Ripple-Down rule learning algorithm, was published in the journal AI Communications. We have continued the development of semantic data mining, which enables the use of ontologies in machine learning. On the one hand, ontologies can be used as background knowledge of the application domain (our paper published in the Journal of Biomedical Informatics presents the analysis of gene expression based on the ontology of gene functions, processes and interactions). On the other hand, an ontology of data mining methods can be used for the automated planning of data-mining workflows in the framework of a novel service-oriented data-mining architecture (three workshop papers were published and a workshop, Service-Oriented Knowledge Discovery, was organized at the ECML/PKDD-08 conference in Antwerp in September 2008). In collaboration with the National Institute of Biology, Ljubljana, Slovenia, we have continued the development of a cost-optimization system for micro-array testing for genetically modified organisms in food and feed samples.

We have developed a model of the Slovenian Health Care network for the Ministry of Health of Slovenia in the scope of the MediNet and MediNet+ projects. In 2008, the project results were upgraded with a web portal that enables querying for analysis results and which offers numerous reporting and visualization facilities. As a result, it can be used to support decision making, monitoring and planning of the Slovenian health-care network.

We successfully concluded the coordination of the FP6 project “Inductive Queries for Mining Patterns and Models” (IQ), which was concerned with developing techniques and applications for inductive databases and queries. Besides data, inductive databases contain generalizations (patterns and models) valid in the data, which are generated, accessed and manipulated by inductive queries. In its final evaluation, the IQ project received much praise, both for its outstanding scientific content and our excellent management. Among our scientific contributions to the project, we wish to highlight the proposed general framework for data mining, which has led to the development of an ontology for data mining. We have also developed many specific approaches to mining complex data, such as ensembles of decision trees for hierarchical multi-label classification, which were applied to

In 2008 the Knowledge Technologies research programme was evaluated as the best research programme in ICT (for the period 2004–2008) by the Slovenian Research Agency (ARRS).

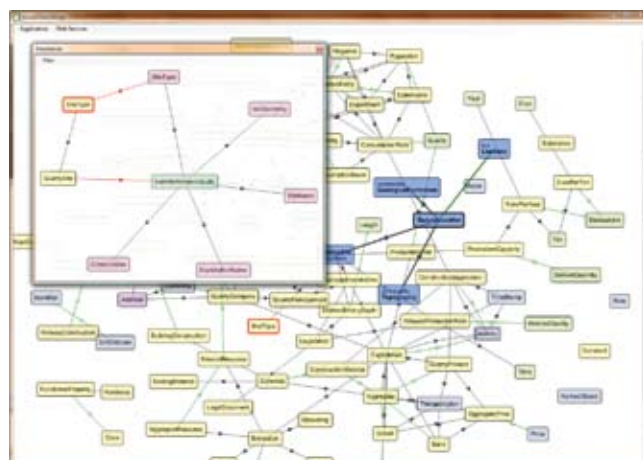
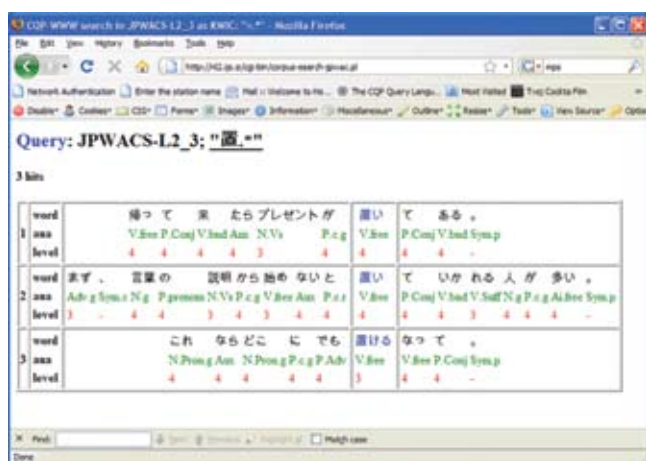


Figure 1: The main window of Visual OntoBridge: the software being developed by our department in the European project SWING. The software provides facilities for annotating web resources with entities from the domain ontology with the purpose of establishing semantic descriptions based on a common vocabulary.



word	part of speech	level	text
国	Noun	1	国は、世界の中心に位置しています。
国	Noun	2	国は、世界の中心に位置しています。
国	Noun	3	国は、世界の中心に位置しています。

Figure 2: An example of concordances for the corpus targeted at learners of Japanese. Available at <http://nl.ijs.si/jaslo/cqp/>.

The successfully completed IQ project, which we coordinated, received much praise from the EU reviewers for outstanding scientific content and excellent management.



Figure 3: Prototype speech synthesiser developed in the project "BMT: development of a text speech reader for mobile telephones for blind and partially sighted users".

In 2008 we started to collaborate in eight new EU projects (3 IPs, 2 STREPs, 1 NoE, 1 CSA) and acquired about 30% of the Slovenian FP7 ICT funds (over €2.5 m), which makes our department one of the most successful in Europe and by far the most successful in Slovenia.

problems in bioinformatics (e.g., functional genomics) and environmental sciences. In addition, the IQ project addressed the problem of using equation discovery for systems biology, which will be further studied in one national and in one FP7 research project that started this year. A national applied research project that deals with the problem of estimating the state of forest ecosystems from remote sensing data was also successfully completed. The forest-vegetation height and the canopy cover are predicted from satellite images. To this end, ensembles of predictive models are used, which are learned from a combination of 2D satellite images and 3D LIDAR data. Maps of forest-vegetation height and canopy cover can now be produced for much larger areas than with previous pilot studies.

In the area of **decision support**, our long-term goal is to develop methods and techniques for decision modelling, support them with software and integrate them with data-mining systems. In 2008 we have improved our software for qualitative multi-criteria decision making, DEXi. We have added three methods for the analysis of decision options: a comparison of options, a selective explanation and a "plus-minus-1 analysis". DEXi 3.0 is freely available from the web page: <http://kt.ijs.si/MarkoBohanec/dexi.html>.

The developed decision-support methods and tools were successfully used in the EU projects SIGMEA, Co-Extra and HEALTHREATS. The first two projects are focused on the analyses of impacts of producing and using genetically modified (GM) plants in Europe: SIGMEA from the viewpoint of ecological and economic impacts, and Co-Extra from the viewpoint of co-existence and traceability in production and supply chains. We concluded our contribution

to SIGMEA with a publication in the journal Ecological Modelling, where we described our model for the assessment of the ecological and economic impacts of growing GM maize. In the ongoing Co-Extra project we have developed three multi-criteria models for the assessment of analytical and sampling methods, as well as for the assessment of the likelihood of the emergence of unapproved GM organisms in food and feed products. In the EU project HEALTHREATS, which aims to develop a decision-support system for the rapid, efficient and coordinated response to threats to health (such as epidemics), we have developed models for an internal evaluation of the work and the results of the project.

In the area of **text and web mining** our work in the FP6 project "Statistical Multilingual Analysis for Retrieval and Translation" (SMART) was focused on testing and using our automatic evaluation system (of machine translation and cross lingual information retrieval) by the project partners, coordination of a user-evaluation study (with different user groups, tasks, datasets and evaluation measures) and an extension of the KCCA method for cross-lingual data analysis in terms of scalability and performance in collaboration with the FP6 project PASCAL. The developed component was integrated into the existing system for contextual search SearchPoint and applied to the problem of cross-lingual information retrieval (demo at SIGIR-2008). As a part of our activities in the FP6 project "Image-based Navigation in Multimedia Archives" (IMAGINATION) we have developed software components to help enrich the existing textual metadata. In particular, the developed software components support three functionalities: (1) annotations of images, (2) discovery of duplicate candidates for ontology concepts and (3) proposing new concepts and relations between them. (1) The annotation of images is based on using textual information that is provided for each image in order to annotate the image with concepts of a predefined ontology. Each image has a very limited amount of text associated with it, usually just a handful of keywords. In some cases additional content is provided via a link to external web pages. Thus, in the proposed approach we handle these two cases separately. (2) The discovery of potential duplicates in an ontology is based on representing each ontology concept by text that can be associated with it, originating either from the concept name and metadata, from external web pages or from the name of an associated URL. (3) The proposing of new ontology concepts and some relations between them is based on using external web pages providing information on particular concepts.

We have successfully concluded two FP6 projects, PASCAL and KDUBiq, where our research in the past year included (1) the development of the visualization of named entities in time based on the use of Wikipedia for extracting named entities and relating them over time, (2) the development of SearchPoint – a system for interfacing an existing search engine that

organizes its results into groups or shows the search results in the context of some existing ontology or classification schema, and (3) the development of Semantic Pointer – a system for social browsing with semantic context that enables real-time identification of the visitors to semantically similar web pages and real-time interaction between them. We have obtained a new FP7 Network of Excellence “Pattern Analysis, Statistical Modelling and Computational Learning 2” (PASCAL2) for the period 2008–2013 where our work will be mainly on research in text mining and natural language processing.

In the area of the **Semantic Web** we have successfully concluded two National Targeted Research projects on “Statistical Semantic Web Systems” (Sistemi za statistični semantični splet) and “Metaservices – Semantic Reasoning Grid Services” (MetaStoritev - Semantično sklapljanje Grid storitev, 2006–2008), where our main contributions are: the development of methods and components for (1) knowledge acquisition, (2) cross-modal data analysis, (3) handling cross-lingual data and several support components for (4) supporting scalability, visualization and evaluation. Our activities in the FP6 project “Lifecycle Support for Networked Ontologies” (NeOn) in the past year resulted in an extension of the NeOn Toolkit with: (1) an alignment server incorporating our pragmatic approach to using large-scale ontologies as contexts (based on a lightweight ontology model and grounding of the ontology concepts in textual documents), (2) a system for ontology visualization in the context of a predefined landscape. Work in the FP6 project “Transitioning Applications to Ontologies” (TAO) resulted in a methodology for cross-modal data analysis focused on connecting textual data and graph/network. We have also designed an evaluation of the methodology on three real-world case studies. In the FP6 project “Semantic Web Services Interoperability for Geospatial Decision Making” (SWING) we have developed the semantic annotation engine OntoBridge and designed several evaluation scenarios on real-world data. We have obtained a new large FP7 IP project “Enabling the Knowledge Powered Enterprise” (ACTIVE) for the period 2008–2011, where our work is mainly in the direction of modelling informal knowledge processes in an organization.

We represent the Jožef Stefan Institute in the World Wide Web Consortium (W3C), which develops and recommends future web standards. We actively participate in the Rule Interchange Format (RIF) working group. The group has so far produced a working draft and last-call recommendations for a number of standards: RIF Core, Basic Logic Dialect, Framework for Logic Dialects, RDF and OWL Compatibility, Datatypes and Built-Ins, Production Rules Dialect, Use Cases and Requirements.

In the area of **knowledge management** we have successfully concluded three FP6 research projects: “European collaborative networked organizations leadership initiative” (ECOLEAD), “Extended Enterprise management in Enlarged Europe” (E4) and “Open Source Enterprise Resource Planning and Order Management System for Eastern European Tool and Die Making Workshops” (Tool-East). The project results have already been transferred to new projects as well as to industrial applications for KOGAST, EMO, the High-Tech cluster, the Toolmakers cluster and the Slovenian automotive cluster. We are continuing the R&D work in the areas of networked organizations, interoperability and collaboration in the FP7 integrated project “Collaboration and Interoperability for networked enterprises” (COIN). We are very successful in the development of Intelligent Cargo Transport in the FP7 project “European Inter-Disciplinary Research on Intelligent Cargo for Efficient, Safe and Environment-friendly Logistics” (EURIDICE), where we have already demonstrated some prototypes and in the FP7 project “Enabling the Knowledge Powered Enterprise” (ACTIVE) with the tools and methods for enterprise “hidden” knowledge formalization.

In the area of **language technologies** we have continued work on the basic research project “Linguistic annotation of Slovene language: methods and resources” (JOS) in the scope of which we are developing automatic inductive methods for the annotation of morphosyntax, syntax and semantics and use these methods to produce freely available linguistically annotated corpora of Slovene.

We finished work on the SEE.ERA-NET project “Building Language Resources and Translation Models for Machine Translation focused on South Slavic and Balkan Languages”, where we produced a parallel and linguistically annotated corpus of EU legal documents, which contains aligned texts in English, Slovene, Serbian, Bulgarian and Romanian.

Our knowledge transfer to industry is demonstrated by the help given to twelve Slovenian companies in becoming partners of EU projects, thus securing approximately €2 m of co-financing.

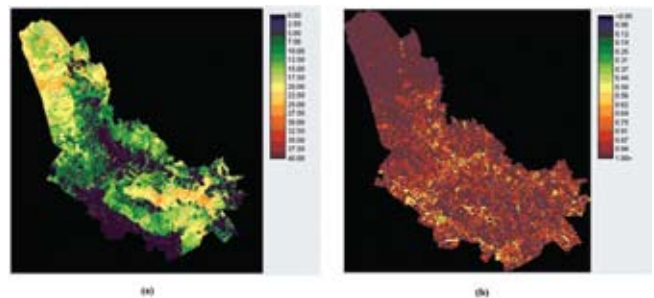


Figure 4: Forest-vegetation height in m (a) and canopy cover in % (b) for the Slivnica region of Slovenia, as predicted from satellite images by ensembles of predictive models learned from a combination of 2D satellite images and 3D LIDAR data.

The Videolectures.net web portal, which we co-manage with the CT3 department, has started to host lectures from the MIT portal “Open Course Ware” and the Opencast community.

The software we developed is used by the EU Commission, the New York Times, Microsoft and many other international companies.

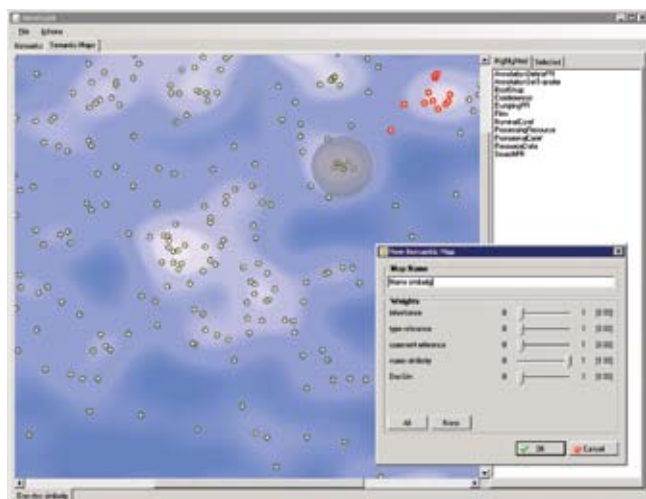


Figure 5: OntoSight: visualization of software components (specifically, Java classes) in the form of a semantic map. This visualization is made by projecting high-dimensional multi-modal data instances onto a 2D canvas by employing multi-dimensional scaling (MDS) and a least-squares solver. The projection preserves distances between instances as much as possible.

We have successfully completed work on a six-month R&D project from the area of eContent and eServices “BMT: development of a text speech reader for mobile telephones for blind and partially sighted users” under the coordination of the company Alpineon R&D, where the JSI was in charge of the development of a morphosyntactic tagger and lemmatiser that improves the quality of the speech synthesis. For Alpineon we also completed the XML_FED project, where we developed the XML specifications for use with a speech synthesiser.

We started work on several new EU projects. In the scope of the project “Conceptual Modelling of Networking of Centres for High-Quality Research in Slavic Lexicography and Their Digital Resources” (MondiLex) and in cooperation with the F-9 department we are investigating the possibilities of using grid technologies for the processing and dissemination of large corpora. In the project “Fostering Language Resources Network” (FlareNet) we are working on the compilation of language resources for Slovene.

In the scope of the bilateral Japanese-Slovenian cooperation with Tokyo International University, we compiled, on the basis of the previously produced jpWaC corpus, a corpus of sentences appropriate for learners of Japanese, where the words are annotated with their difficulty level. The corpus is freely available for use via a web concordancer.

Some outstanding publications in the past three years

1. John Davies (Ed.), Marko Grobelnik (Ed.), Dunja Mladenić (Ed.). Semantic Knowledge Management: Integrating Ontology Management, Knowledge Discovery, and Human Language Technologies. Springer, 2008.
2. Igor Trajkovski, Nada Lavrač, Jakub Tolar. SEGS: search for enriched gene sets in microarray data. Journal of biomedical informatics, 2008, vol. 41, no. 4, pp. 588-601.
3. Gemma C. Garriga, Petra Kralj Novak, Nada Lavrač. Closed sets for labeled data. Journal of machine learning research, 2008, vol. 9, no. 4, pp. 559-580.
4. Sašo Džeroski, Ljupčo Todorovski. Equation discovery for systems biology: finding the structure and dynamics of biological networks from time course data. Current opinion in biotechnology, 2008, 19: 360-368.
5. Marko Bohanec, Antoine Messean, Sara Scatasta, Frédérique Angevin, Bryan Griffiths, Paul Henning Krogh, Martin Žnidaršič, Sašo Džeroski. A qualitative multi-attribute model for economic and ecological assessment of genetically modified crops. Ecol. model., 2008, vol. 215, no. 1/3, pp. 247-261.

Awards and appointments

1. Sašo Džeroski, ECCAI fellow – ECCAI award. Awarded by European Coordination Committee for Artificial Intelligence

Organization of conferences, congress and meetings

1. Information Society 2008, organization of subconferences: SiKDD-2007, Intelligent Systems, 13-17 Oct. 2008

INTERNATIONAL PROJECTS

1. Systems Biology of Phagosome Formation and Maturation, Modulation by Intracellular Pathogens
PHAGOSYS
7. FP, 223451, HEALTH-F4-2008-223451
EC; Anne-Marie Fish, Imperial College of Science, Technology and Medicine, London, Great Britain
Prof. Sašo Džeroski
2. Biosocial Networks for Creative Information Discovery
BISON
7. FP, 211898
EC; Universität Konstanz, Konstanz, Germany
Prof. Nada Lavrač
3. Conceptual Modelling of Networking of Centres for High-Quality Research in Slavic Lexicography and Their Digital Resources
MONDILEX
7. FP, 211938

EC; Institute of Mathematics and Informatics of the Bulgarian Academy of Science, Sofia, Bulgaria

Asst. Prof. Tomaž Erjavec, Jan Jona Javoršek, B. Sc.

4. Enabling the Knowledge Powered Enterprise

ACTIVE

7. FP, 215040

EC; Philip Hewitt, British Telecommunications plc, London, Great Britain

Marko Grobelnik, B. Sc., Asst. Prof. Dunja Mladenić, Mitja Jermol, M. Sc.

5. Pattern Analysis, Statistical Modelling and Computational Learning 2

PASCAL 2

7. FP, 216886

EC; Eileen Simon, University of Southampton, Highfield, Southampton, Great Britain

Asst. Prof. Dunja Mladenić, Marko Grobelnik, B. Sc., Mitja Jermol, M. Sc.

6. European Inter-Disciplinary Research on Intelligent Cargo for Efficient, Safe and

Environment-friendly Logistics

EURIDICE

7. FP, 216271

EC; INSIEL - Informatica per il Sistema degli Enti Locali s.p.a, Trieste, Italy

Marko Grobelnik, B. Sc., Asst. Prof. Dunja Mladenić, Mitja Jermol, M. Sc.

7. Collaboration and INteroperability for networked enterprises
COIN
7. FP, 216256
EC; Claudia Guglielmina, TXT e-Solutions Spa, Milano, Italy
Asst. Prof. Dunja Mladenić, Marko Grobelnik, B. Sc., Mitja Jermol, M. Sc.
8. GM and non-GM supply chains: their CO-Existence and TRAceability
Co-Extra
6. FP, 007158
EC; Institut National de la Recherche Agronomique, Paris, France
Prof. Marko Bohanec
9. European Embryonal Tumor Pipeline
E.E.T. -Pipeline
6. FP, 037260
EC; Angelika Eggert, Universitaet Duisburg-Essen, Essen, Germany
Prof. Sašo Džeroski
10. Stimulating Policy Debate on Women and Science Issues in Central Europe
WS DEBATE
6. FP, 036651
EC; Dr. Dora Groo, Eszter Papp, Hungarian Science and Technology Foundation; Tudományos és Technológiai Alapítvány, Budapest, Hungary
Asst. Prof. Dunja Mladenić, Marko Grobelnik, B. Sc., Mitja Jermol, M. Sc.
11. Statistical Multilingual Analysis for Retrieval and Translation
SMART
6. FP, 033917
EC; Nicola Cancedda, Xerox Research Centre Europe, Meylan; Xerox, Aulnay-Sous-Bois, France
Asst. Prof. Dunja Mladenić, Marko Grobelnik, B. Sc., Mitja Jermol, M. Sc.
12. Image-based Navigation in Multimedia Archives
IMAGINATION
6. FP, 034626
EC; Clemens van Dinter, Forschungszentrum Informatik an der Universität Karlsruhe, Karlsruhe, Germany
Asst. Prof. Dunja Mladenić, Mitja Jermol, M. Sc., Prof. Marko Mikuz
13. Extended Enterprise Management in Enlarged Europe
E4
6. FP, 027282
EC; Roberto Tarditi, Centro Ricerche Fiat Societa Consortile per Azioni, Orbassano (TO), Italy
Asst. Prof. Dunja Mladenić, Marko Grobelnik, B. Sc., Mitja Jermol, M. Sc.
14. Open Source Enterprise Resource Planning and Order Management System for Eastern European Tool and Die Making Workshop
Tool-East
6. FP, 027802
EC; dr.-ing. Volker Stich, Forschungsinstitut fuer Rationalisierung (FIR) and der RWTH Aachen, Research Institute for Operations Management at Aachen University, Aachen, Germany
Asst. Prof. Dunja Mladenić, Marko Grobelnik, B. Sc., Mitja Jermol, M. Sc.
15. Semantic Web Services Interoperability for Geospatial Decision Making
SWING
6. FP, 026514
EC; Arne J. Berre, SINTEF - Stiftelsen for Industriell OG Teknisk Forskning Ved Norges Tekniske Hogskole, Trondheim; SINTEF ICT, Oslo, Norway
Asst. Prof. Dunja Mladenić, Marko Grobelnik, B. Sc., Mitja Jermol, M. Sc., Prof. Marko Mikuz
16. Lifecycle Support for Networked Ontologies
NEON
6. FP, 027595
EC; Prof. Enrico Motta, Kmi, The Open University, Milton Keynes, Great Britain
Asst. Prof. Dunja Mladenić, Marko Grobelnik, B. Sc., Mitja Jermol, M. Sc.
17. Transitioning Applications to Ontologies
TAO
6. FP, 026460
EC; Dr. Kalina Bontcheva, University of Sheffield, Department of Computer Science, Sheffield, Great Britain
Asst. Prof. Dunja Mladenić, Marko Grobelnik, B. Sc., Mitja Jermol, M. Sc.
18. Inductive Queries for Mining Patterns and Models
IQ
6. FP, 516169
EC; Prof. Sašo Džeroski, Institut "Jožef Stefan", Ljubljana, Slovenia
Prof. Sašo Džeroski
19. European Collaborative networked Organizations LEADership initiative
ECOLEAD
6. FP, 506958
EC; Martin Ollus, Technical Research Centre of Finland, Espoo, Finland
Prof. Nada Lavrač, Mitja Jermol, M. Sc.
20. Pattern Analysis, Statistical Modelling and Computational Learning
PASCAL
6. FP, 506778
EC; Prof. John Shawe-Taylor, The University of Southampton, School of Electronics and Computer Science, Highfield, Southampton, Great Britain
Asst. Prof. Dunja Mladenić, Mitja Jermol, M. Sc.
21. KD-ubiq - A Blueprint for Ubiquitous Knowledge Discovery Systems
KD-ubiq
6. FP, 021321
EC; Dr. Michael May, Stephan Kollmer, Fabian Perpeet, Fraunhofer Gesellschaft zur Förderung der Angewandten Forschung e.V., München; Sankt Augustin, Germany
Asst. Prof. Dunja Mladenić
22. Integrated Decision Support System for HEALTH THREATS and Crises Management
HEALTHTHREATS
Public Health program (PHEA), e-Contentplus
2006203
EC; Executive Agency for Public Health (PHEA - HTC), Luxembourg; Azienda Sanitaria Locale di Brescia (ASL Brescia), Brescia, Italy
Prof. Nada Lavrač, Dr. Martin Žnidaršič
23. Fostering Language Resources Network
FLaReNet
e-Contentplus
ECP-2007-LANG-617001
EC; CNR-ILC, Consiglio Nazionale delle Ricerche, Rome, Italy
Asst. Prof. Tomaž Erjavec
24. Building Language Resources and Translation Models for Machine Translation focused on South Slavic and Balkan Languages
SEE-ERA.NET
1000-07-380024
Radoslav Pavlov, Institute of Mathematics and Informatics, Bulgarian Academy of Sciences, Sofia, Bulgaria
Asst. Prof. Tomaž Erjavec
25. The Use of a Dedicated Service on the Http Server of IJS <http://nl.ijs.si> to be used by FP for Up-loading and Storing Texts which Constitute the FP Corpus of IX Century Translated Books
Forschungsprojekt: Deutsch-slowenische/kroatische Übersetzung 1848 bis 1918
Agreement, dated 3.5.2007
Prof. Erich Prunč, Graz, Austria
Asst. Prof. Tomaž Erjavec
26. Information Fusion for Data Mining in Bioinformatics and Biomedicine
BI-CZ/08-09-007
Prof. Olga Štěpánková, Czech Technical University in Prague, Prague, Czech Republic
Prof. Nada Lavrač
27. Inductive Databases for Genomics and Proteomics
BI-HR/07-08-029
Dr. Tomislav Šmuc, Rudjer Boskovic Institute, Zagreb, Croatia
Prof. Sašo Džeroski
28. Japanese/Slovene Resources for Second Language Learners
BI-JP/08-10/006
Prof. Yoshiko Kawamura, Tokyo International University, Saitama, Japan
Asst. Prof. Tomaž Erjavec
29. Knowledge Discovery for Ecological Modeling of Lake Ecosystems
BI-MK/07-08-017
Prof. Kosta Mitreski, Faculty of Electrical Engineering, Skopje, The Republic of Macedonia
Prof. Sašo Džeroski
30. Landscape Functioning under Different Management Regimes
LANDECO
SCRI, Living Technology, Dundee, Scotland, Great Britain
Asst. Prof. Marko Debeljak
31. Analysis of Dynamic Networks with Graph and Text Mining Methods
BI-US/06-07-032
Faloutsos Christos, Carnegie Mellon University, Pittsburgh, PA, USA
Asst. Prof. Dunja Mladenić

R & D GRANTS AND CONTRACTS

1. Advanced ML methods for automated modeling of dynamic systems
Prof. Sašo Džeroski
2. Linguistic annotation of Slovene language: methods and resources
Asst. Prof. Tomaž Erjavec
3. Methodological aspects of cognitive process research-learning and decision-making
Prof. Andrej Ule, Prof. Marko Bohanec
4. Unknown 17th and 18th century manuscripts of Slovenian literature: information-technology aided register, scholarly editions and analyses
Dr. Matija Ogrin, Asst. Prof. Tomaž Erjavec
5. Slovene Terminology Web Portal
Prof. Vojko Gorjanc, Simon Krek
6. Methodology for producing a detailed digital map of the height and density of vegetation cover
Prof. Sašo Džeroski
7. VoiceTRAN II - Multilingual mobile speech communicator for 21.th century warriors
Dr. Jerneja Zganec Gros, Asst. Prof. Tomaž Erjavec
8. Crisis management simulator
Dr. Matej Penca, Asst. Prof. Dunja Mladenić
9. Development of Knowledge Management System for SV
Asst. Prof. Jaroslav Berce, Marko Grobelnik
10. Statistical semantic web systems
Asst. Prof. Dunja Mladenić
11. Metaservices - Semantic reasoning Grid services
Dr. Gregor Pipan, Asst. Prof. Dunja Mladenić
12. Digital text centre with multimedia communication
Dr. Matija Ogrin, Asst. Prof. Tomaž Erjavec

13. Harmonisation of technologies for following genetically modified organisms in food and feed production chain and its co-existence with conventional and ecological production chains
Asst. Prof. Vladimir Meglič, Prof. Nada Lavrač
14. Guidelines for national strategy of preservation of forest trees genefond due to introduction of genetically modified organisms in agriculture
Dr. Robert Brus, Prof. Marko Debeljak
15. Harmful factors for contemporary forests: methods for monitoring and ecological modelling, the impact of exploitation, and strategies for management
Prof. Maja Jurc, Prof. Sašo Džeroski
16. Analysis and scenario of development and exploration of forests in Slovenia
Dr. Andrej Bončina, Prof. Marko Debeljak
17. Influence of gene transfer, genetic diversity and cultivation technology on sustainability and tracing methods
Asst. Prof. Vladimir Meglič, Prof. Marko Debeljak
18. European Uniform System for Linking Digital Information of Public Nature: DUNE
Marko Grobelnik

19. Development of a text speech reader for mobile telephones for blind and partially sighted users: BMT
Asst. Prof. Tomaž Erjavec

RESEARCH PROGRAM

1. Knowledge Technologies
Prof. Nada Lavrač

NEW CONTRACT

1. Systems Biology Tools Development for Cell Therapy and Drug Development - SYSTHER
National Institute of Biology
Prof. Nada Lavrač, Petra Kralj Novak

VISITORS FROM ABROAD

1. Prof. Luis Torgo, Faculty of Economics, University of Porto, Portugal, 5-8 Jan. 2008 and 10-14 Dec. 2008
2. Jeroen Sebastian de Bruin, Leiden University of Advanced Computer Science, Leiden, The Netherlands, 8-20 Jan. 2008
3. Tricia Jenkins, University of Liverpool, Liverpool, United Kingdom, 22-23 Jan. 2008
4. Margaret Evans, University of Liverpool, Liverpool, United Kingdom, 22-23 Jan. 2008
5. Prof. Joost Kok, Leiden University, Leiden, The Netherlands, 21-24 Jan. 2008
6. Prof. John Shawe Taylor, School of Electronics and Computer Science, University of Southampton, Southampton, United Kingdom, 31 Jan. 2008
7. Dr Hendrich Blockeel, Katholieke Universiteit Leuven, Belgium, 4-8 Feb. 2008
8. Prof. Suzana Loskovska, Faculty of Electrical Engineering and Information, Skopje, Macedonia, 15-22 Feb. 2008
9. dr. John Davies, British Telecom, Ipswich, United Kingdom, 16. Feb. 2008
10. Delia Rusu, Technical University of Cluj-Napoca, Cluj-Napoca, Romania, 18 Feb. to 7 Jun. 2008
11. Lorand Dali, Technical University of Cluj-Napoca, Cluj-Napoca, Romania, 18 Feb. to 7 Jun. 2008
12. Christophe Schwartz, ENSAIA/INRA Nancy, France, 25-29 Feb. 2008
13. Jerome Cortet, ENSAIA/INRA Nancy, France, 25-29 Feb. and 6-10 Oct. 2008
14. Rayid Ghani, Accenture Labs, Chicago, USA, 14 Mar. 2008
15. Elena Mitreska, Faculty of Electrical Engineering and Information, Skopje, Macedonia, 16-23 Mar. 2008
16. doc. dr. Kosta Mitreski, Faculty of Electrical Engineering and Information, Skopje, Macedonia, 16-23 Mar. 2008
17. Andreja Naumoski, Faculty of Electrical Engineering and Information, Skopje, Macedonia, 16-23 Mar. 2008
18. dr. Paul B. Losiewicz, European Office of Aerospace Research and Development, United Kingdom, 21-22 Apr. 2008
19. Florence Leprince, ARVALIS-Institut du végétal 21 chemin de Pau, Montradon, France, 22 Apr. 2008
20. dr. Neel Sundaresan, eBay Research, San Jose, USA, 15 May 2008
21. Walter Scholger, Karl-Franzens-Universität, Graz, Austria, 18 May 2008
22. Brita Luber, Karl-Franzens-Universität, Graz, Austria, 18 May 2008
23. Petra Schikler, Karl-Franzens-Universität, Graz, Austria, 18 May 2008
24. Dr Hubert Stigler, Department of Information Processing in the Humanities, Karl-Franzens-Universität, Graz, Austria, 19 May 2008
25. Dr Nataša Milić-Frayling, Microsoft Research, Cambridge, United Kingdom: 11-12 Jun. 2008
26. r. Arno Knobbe, Utrecht University, Utrecht, The Netherlands: Exceptional Model Mining, 17 Jun. 2008
27. dr. Michael Witbrock, Cycorp, Inc., Austin, USA, 26 Jun. to 6 Sept. 2008
28. Alexandra Moraru, Technical University of Cluj-Napoca, Cluj-Napoca, Romania, 7 Jul. 15 Sept. 2008
29. Titus Moldovan, Technical University of Cluj-Napoca, Cluj-Napoca, Romania, 7 Jul. 15 Sept. 2008
30. Paul-Valentin Borza, Technical University of Cluj-Napoca, Cluj-Napoca, Romania, 13 Jul. to 21 Sept. 2008
31. dr. Hubert Stigler s Karl-Franzens-Universität, Graz, Austria, 25 Jul. 2008 and 18 May 2008
32. Dr Stefano Bertolo, European Commission, Brussels, Belgium, 19 Aug. 2008
33. Mandarić Domagoj, FirstData International, Zagreb, Croatia, 14 Aug. 2008
34. Timo Aho, Tampere University of Technology, Tampere, Finland, 1 Oct. 2008 to 31 Mar. 2009
35. Monika Žakova, Czech Technical University Prague, Prague, Czech Republic, 13-24 Oct. 2008
36. Dr John Bullas, Atkins Highways & Transportation, London, United Kingdom, 16-19 Oct. 2008
37. Mirjana Ivanović, Faculty of Science University of Novi Sad, Novi Sad, Serbia, 16-19 Oct. 2008
38. Dr Tomislav Šmuc, Institut Rudjer Bošković, Zagreb, Croatia, 21 Nov. 2008
39. dr. Dragan Gamberger, Institut Rudjer Bošković, Zagreb, Croatia, 21 Nov. 2008
40. Mislav Malenica, Institut Rudjer Bošković, Zagreb, Croatia, 21 Nov. 2008
41. Ivica Dimitrovski, Faculty of Electrical Engineering and Information, Skopje, Macedonia, 28 Sept. to 18 Oct. 2008

STAFF

Researchers

1. Prof. Marko Bohanec
2. Dr. Damjan Bojadžiev
3. Dr. Bojan Cestnik*
4. Prof. Marko Debeljak
5. Prof. Sašo Džeroski
6. Asst. Prof. Tomaž Erjavec
7. **Prof. Nada Lavrač, Head**
8. Asst. Prof. Dunja Mladenić
9. Prof. Tanja Urbančič*

Postdoctoral associates

10. Dr. Branko Kavšek*
11. Dr. Bernard Ženko
12. Dr. Martin Žnidaršič

Postgraduates

13. Janez Brank, M. Sc.
14. Darko Čerepnalkoski, B. Sc.
15. Jure Ferlež, M. Sc.
16. Blaž Fortuna, B. Sc.
17. Valentin Gjorgjioski, B. Sc.
18. Miha Grčar, B. Sc.
19. Mitja Jermol, M. Sc.
20. Dragi Kocev, B. Sc.

21. Petra Kralj Novak, B. Sc.
22. Simon Krek*, B. Sc.
23. Blaž Novak
24. Panče Panov, B. Sc.
25. Vid Podpečan, B. Sc.
26. Jan Rupnik, B. Sc.
27. Mitja Trampuš, B. Sc.
28. Dr. Miha Volovšek*
29. Miha Vuk, B. Sc.

Technical officers

30. Milica Bauer, B. Sc.
31. Dr. France Dacar
32. Dr. Damjan Demšar
33. Dr. Igor Mozetič
34. Nina Novinec, M. Sc.

Technical and administrative staff

35. Tina Anžič
36. Marko Grobelnik
37. Jolanda Jakofčič
38. Boštjan Pajntar

Note:

* part-time JSI member

BIBLIOGRAPHY

ORIGINAL ARTICLES

1. Nataša Atanasova, Ljupčo Todorovski, Sašo Džeroski, Boris Kompare, "Application of automated model discovery from data and expert knowledge to a real-world domain: Lake Glumsø", In: *Proceedings of The Fifth European Conference on Ecological Modelling ECEM 2005, Pushchino, Russia, September 19-23*, (Ecological modelling, Vol. 212, no. 1/2, 2008), Alexander S. Komarov, ed., Amsterdam, Elsevier Scientific Publ. Co., 2008, vol. 212, no. 1/2, pp. 92-98, 2008.
2. Marko Bohanec, Antoine Messean, Sara Scatasta, Frédérique Angevin, Bryan Griffiths, Paul Henning Krogh, Martin Žnidaršič, Sašo Džeroski, "A qualitative multi-attribute model for economic and ecological assessment of genetically modified crops", In: *Proceedings of ICEM 2006, International Conference on Ecological Modelling, Mamaguchi, Japan: 28 August - 1 September 2006, Yamaguchi University, Japan*, (Ecological modeling, Vol. 215, no. 1/3, 2008), Amsterdam, Elsevier Scientific Publ. Co., 2008, vol. 215, no. 1/3, pp. 247-261, 2008.
3. Janez Brank, Dunja Mladenić, Marko Grobelnik, Nataša Milić-Frayling, "Feature selection for the classification of large document collections", *J. univers. comput. sci. (Online)*, vol. 14, no. 10, pp. 1562-1596, 2008.
4. Marko Debeljak, Geoff Squire, Damjan Demšar, Marc W. Young, Sašo Džeroski, "Relations between the oilseed rape volunteer seedbank, and soil factors, weed functional groups and geographical location in the UK", In: *Proceedings of The Fifth European Conference on Ecological Modelling ECEM 2005, Pushchino, Russia, September 19-23*, (Ecological modelling, Vol. 212, no. 1/2, 2008), Alexander S. Komarov, ed., Amsterdam, Elsevier Scientific Publ. Co., 2008, vol. 212, no. 1/2, pp. 138-146, 2008.
5. Sašo Džeroski, Ljupčo Todorovski, "Equation discovery for systems biology: finding the structure and dynamics of biological networks from time course data", *Curr. opin. biotechnol.*, vol. 19, no. 4, pp. 360-368, 2008.
6. Blaž Fortuna, Nada Lavrač, Paola Velardi, "Advancing topic ontology learning through term extraction", In: *PRICAI 2008: trends in artificial intelligence: proceedings*, (Lecture notes in computer science, Lecture notes in artificial intelligence, vol. 5351), 10th Pacific Rim International Conference on Artificial Intelligence, Hanoi, Vietnam, December 15-19, 2008, Tu-Bao Ho, ed., Berlin, Heidelberg, Springer, 2008, vol. 5351, pp. 626-635, 2008.
7. Dragan Gamberger, Nada Lavrač, Johannes Fürnkranz, "Handling unknown and imprecise attribute values in propositional rule learning: a feature-based approach", In: *PRICAI 2008: trends in artificial intelligence: proceedings*, (Lecture notes in computer science, Lecture notes in artificial intelligence, vol. 5351), 10th Pacific Rim International Conference on Artificial Intelligence, Hanoi, Vietnam, December 15-19, 2008, Tu-Bao Ho, ed., Berlin, Heidelberg, Springer, 2008, vol. 5351, pp. 636-645, 2008.
8. Gemma C. Garriga, Petra Kralj Novak, Nada Lavrač, "Closed sets for labeled data", *J. mach. learn. res.*, vol. 9, no. 4, pp. 559-580, 2008.
9. Marko Grobelnik, Janez Brank, Blaž Fortuna, Igor Mozetič, "Contextualizing ontologies with ontolight: a pragmatic approach", *Informatica (Ljublj.)*, vol. 32, no. 1, pp. 79-84, 2008.
10. Kristina Hmeljak Sangawa, Tomaž Erjavec, "A low cost approach to building a Japanese-Slovene parallel corpus", *Denshi Jōhō Tsūshin Gakkai giyutsu kenkyū hōkokoku*, vol. 108, no. 50, pp. 7-10, 2008.
11. Aneta Ivanovska, Celine Vens, Nathalie Colbach, Marko Debeljak, Sašo Džeroski, "The feasibility of co-existence between conventional and genetically modified crops: using machine learning to analyse the output of simulation models", In: *Proceedings of ICEM 2006, International Conference on Ecological Modelling, Mamaguchi, Japan: 28 August - 1 September 2006, Yamaguchi University, Japan*, (Ecological modeling, Vol. 215, no. 1/3, 2008), Amsterdam, Elsevier Scientific Publ. Co., 2008, issues 1-3, vol. 215, pp. 262-271, 2008.
12. Simon Krek, "FrameNet in slovenščina", *Jez. slovst. (Tisk. izd.)*, vol. 53, no. 5, pp. [37]-54, sep.-okt. 2008.
13. Jana Laganis, Aleksandar Pečkov, Marko Debeljak, "Modeling radial growth increment of black alder (*Alnus glutinosa* (L.) Gaertn.) tree", In: *Proceedings of ICEM 2006, International Conference on Ecological Modelling, Mamaguchi, Japan: 28 August - 1 September 2006, Yamaguchi University, Japan*, (Ecological modeling, Vol. 215, no. 1/3, 2008), Amsterdam, Elsevier Scientific Publ. Co., 2008, issues 1-3, vol. 215, pp. 180-189, 2008.
14. Nada Lavrač, Domen Jesenovec, Nejc Trdin, Neža Mramor Kosta, "Mining spatio-temporal data of traffic accidents and spatial pattern visualization", *Metodol. zv. (Tisk. izd.)*, vol. 5, no. 1, pp. 45-63, 2008.
15. Aleksandar Pečkov, Sašo Džeroski, Ljupčo Todorovski, "A minimal description length scheme for polynomial regression", In: *Advances in knowledge discovery and data mining: 12th Pacific-Asia Conference, PAKDD 2008, Osaka, Japan, May 20-23, 2008: proceedings*, (Lecture notes in computer science, Lecture notes in artificial intelligence, vol. 5012), Takashi Washio, ed., Berlin, New York, Springer, cop. 2008, vol. 5012, pp. 284-295, 2008.
16. Ingrid Petrič, Tanja Urbančič, Bojan Cestnik, Marta Macedoni-Lukšič, "Literature mining method Rajolink for uncovering relations between biomedical concepts", *Journal of biomedical informatics*, 9 pp., Avg. 2008.
17. Sandrine Pivard, Damjan Demšar, Jane Lecomte, Marko Debeljak, Sašo Džeroski, "Characterizing the presence of oilseed rape feral populations on field margins using machine learning", In: *Proceedings of The Fifth European Conference on Ecological Modelling ECEM 2005, Pushchino, Russia, September 19-23*, (Ecological modelling, Vol. 212, no. 1/2, 2008), Alexander S. Komarov, ed., Amsterdam, Elsevier Scientific Publ. Co., 2008, vol. 212, no. 1/2, pp. 147-154, 2008.
18. Joël Plisson, Nada Lavrač, Dunja Mladenić, Tomaž Erjavec, "Ripple down rule learning for automated word lemmatisation", *AI commun.*, vol. 21, no. 1, pp. 15-26, 2008.
19. Irena Srdanović Erjavec, Tomaž Erjavec, Adam Kilgarriř, "A web corpus and word sketches for Japanese", *Shizen gengo shori*, vol. 15, no. 2, pp. 137-159, 2008.
20. Igor Trajkovski, Nada Lavrač, Jakub Tolar, "SEGS: search for enriched gene sets in microarray data", *Journal of biomedical informatics*, vol. 41, no. 4, pp. 588-601, 2008.
21. Igor Trajkovski, Filip Železný, Nada Lavrač, Jakub Tolar, "Learning relational descriptions of differentially expressed gene groups", *IEEE trans. syst. man cybern., Part C Appl. rev.*, vol. 38, no. 1, spec. issue, pp. 16-25, 2008.
22. Veronika Velenšek Prestor, Uroš Mazić, Ciril Kržišnik, Damjan Demšar, Janez Jazbec, Berta Jereb, "Cardiac damage after treatment of childhood cancer: a long-term follow-up", *BMC Cancer*, vol. 8, pp. 141-141-8, 2008.
23. Celine Vens, Jan Struyf, Leander Schietgat, Sašo Džeroski, Hendrik Blockeel, "Decision trees for hierarchical multi-label classification", *Mach. learn.*, vol. 73, no. 2, pp. 185-214, 2008.
24. Bernard Ženko, "Learning predictive clustering rules", *Informatica (Ljublj.)*, vol. 32, no. 1, pp. 95-96, 2008.
25. Bernard Ženko, Sašo Džeroski, "Learning classification rules for multiple target attributes", In: *Advances in knowledge discovery and data mining: 12th Pacific-Asia Conference, PAKDD 2008, Osaka, Japan, May 20-23, 2008: proceedings*, (Lecture notes in computer science, Lecture notes in artificial intelligence, vol. 5012), Takashi Washio, ed., Berlin, New York, Springer, cop. 2008, vol. 5012, pp. 454-465, 2008.

REVIEW ARTICLES AND CHAPTERS IN BOOKS

1. Luis Camarinha-Matos, Ana Ines Oliveira, Damjan Demšar, Michele Sesana, Arturo Molina, Fabiano Baldo, Toni Jarimo, "VO creation assistance services", In: *Methods and tools for collaborative networked organizations*, Luis Camarinha-Matos, ed., Hamideh Afsarmanesh, ed., Martin Ollus, ed., New York, Springer, 2008, pp. 155-189.
2. Sašo Džeroski, "Data mining", In: *Encyclopedia of ecology*, Sven Erik Jørgensen, ed., Amsterdam, London, Elsevier, 2008, zv. 2, pp. 821-30.
3. Toni Jarimo, Peter Ljubić, Jiří Hodík, Marko Bohanec, Nada Lavrač, "Multi-criteria partner selection in virtual organizations", In: *Encyclopedia of networked and virtual organizations*, Goran Putnik, ed., Hershey (PA) [etc.], Information Science Reference, cop. 2008, pp. 964-970.
4. Dunja Mladenić, Nada Lavrač, "A European virtual enterprise on collaborative data mining and decision support", In: *Encyclopedia of networked and virtual organizations*, Goran Putnik, ed., Hershey (PA) [etc.], Information Science Reference, cop. 2008, pp. 524-529.
5. Joël Plisson, Peter Ljubić, Igor Mozetič, Nada Lavrač, "Ontologies for collaborative networked organizations", In: *Encyclopedia of networked and virtual organizations*, Goran Putnik, ed., Hershey (PA) [etc.], Information Science Reference, cop. 2008, pp. 1128-1135.

6. Martin Žnidaršič, Marko Bohanec, Blaž Zupan, "Data-driven revision of decision models", In: *Encyclopedia of data warehousing and mining. Volume II*, John Wang, ed., 2nd ed., Hershey, New York, Information Science Reference, 2008, pp. 617-623.

PUBLISHED CONFERENCE PAPERS

Invited Papers

1. Sašo Džeroski, "Inductive process modeling for systems biology", In: *European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases, EMM/ PKDD 2008: proceedings of Induction of process models, IPM'08, 15 September, 2008, Antwerp, Belgium*, Will Bridewell, ed., Toon Calder, ed., Ana Karla de Medeiros, ed., Stefan Kramer, ed., Mykola Pechenizkiy, ed., Ljupčo Todorovski, ed., [S. l., s. n.], 2008, pp. 3-4.

Regular papers

1. Darko Aleksovski, Martin Erwing, Sašo Džeroski, "A functional programming approach to distance-based machine learning", In: *Zbornik 11. mednarodne multikonference Informacijska družba - IS 2008, 13.-17. oktober 2008: zvezek A: volume A*, (Informacijska družba), Marko Bohanec, ed., Matjaž Gams, ed., Vladislav Rajkovič, ed., Tanja Urbančič, ed., Mojca Bernik, ed., Dunja Mladenčič, ed., Marko Grobelnik, ed., Marjan Heričko, ed., Urban Kordeš, ed., Olga Markič, ed., Ljubljana, Institut "Jožef Stefan", 2008, pp. 161-165.
2. Marko Bohanec, "Učenje in odločanje: analiza definicij osnovnih konceptov v Wikipediji z metodo analize besedil in omrežij", In: *Zbornik 11. mednarodne multikonference Informacijska družba - IS 2008, 13.-17. oktober 2008: zvezek A: volume A*, (Informacijska družba), Marko Bohanec, ed., Matjaž Gams, ed., Vladislav Rajkovič, ed., Tanja Urbančič, ed., Mojca Bernik, ed., Dunja Mladenčič, ed., Marko Grobelnik, ed., Marjan Heričko, ed., Urban Kordeš, ed., Olga Markič, ed., Ljubljana, Institut "Jožef Stefan", 2008, pp. 305-308.
3. Marko Bohanec, Martin Žnidaršič, "Supporting decisions about the introduction of genetically modified crops", In: *Collaborative decision making: perspectives and challenges*, (Frontiers in artificial intelligence and applications, vol. 176), Pascale Zaraté, ed., Amsterdam ... [etc.], IOS Press, 2008, pp. 404-415.
4. Damjan Bojadžiev, "Introspekcija in egzozpekcija", In: *Zbornik 11. mednarodne multikonference Informacijska družba - IS 2008, 13.-17. oktober 2008: zvezek A: volume A*, (Informacijska družba), Marko Bohanec, ed., Matjaž Gams, ed., Vladislav Rajkovič, ed., Tanja Urbančič, ed., Mojca Bernik, ed., Dunja Mladenčič, ed., Marko Grobelnik, ed., Marjan Heričko, ed., Urban Kordeš, ed., Olga Markič, ed., Ljubljana, Institut "Jožef Stefan", 2008, pp. 309-310.
5. Jeroen de Bruin, Joost N. Kok, Nada Lavrač, Igor Trajkovski, "On the design of knowledge discovery services design patterns and their application in a use case implementation", In: *Third generation data mining: towards service-oriented knowledge discovery*, European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases, EMM/ PKDD 2008, 15-19 September, 2008, Antwerp, Belgium, Nada Lavrač, ed., Joost N. Kok, ed., Jeroen de Bruin, ed., Vid Podpečan, ed., [S. l., s. n.], 2008, pp. 43-57.
6. Jeroen de Bruin, Joost N. Kok, Nada Lavrač, Igor Trajkovski, "Towards service-oriented knowledge discovery: a case study", In: *Third generation data mining: towards service-oriented knowledge discovery*, European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases, EMM/ PKDD 2008, 15-19 September, 2008, Antwerp, Belgium, Nada Lavrač, ed., Joost N. Kok, ed., Jeroen de Bruin, ed., Vid Podpečan, ed., [S. l., s. n.], 2008, pp. 1-10.
7. Darko Črepnalkoski, Katerina Taškova, Ljupčo Todorovski, Sašo Džeroski, "Inducing process-based models of dynamic systems from multiple data sets", In: *European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases, EMM/ PKDD 2008: proceedings of Induction of process models, IPM'08, 15 September, 2008, Antwerp, Belgium*, Will Bridewell, ed., Toon Calder, ed., Ana Karla de Medeiros, ed., Stefan Kramer, ed., Mykola Pechenizkiy, ed., Ljupčo Todorovski, ed., [S. l., s. n.], 2008, pp. 5-12.
8. Lorand Dali, Blaž Fortuna, "Triplet extraction from sentences using SVM", In: *Zbornik 11. mednarodne multikonference Informacijska družba - IS 2008, 13.-17. oktober 2008: zvezek A: volume A*, (Informacijska družba), Marko Bohanec, ed., Matjaž Gams, ed., Vladislav Rajkovič, ed., Tanja Urbančič, ed., Mojca Bernik, ed., Dunja Mladenčič, ed., Marko Grobelnik, ed., Marjan Heričko, ed., Urban Kordeš, ed., Olga Markič, ed., Ljubljana, Institut "Jožef Stefan", 2008, pp. 170-173.
9. Ivica Dimitrovski, Dragi Kocev, Suzana Loškovska, Sašo Džeroski, "Hierarchical annotation of medical images", In: *Zbornik 11. mednarodne multikonference Informacijska družba - IS 2008, 13.-17. oktober 2008: zvezek A: volume A*, (Informacijska družba), Marko Bohanec, ed., Matjaž Gams, ed., Vladislav Rajkovič, ed., Tanja Urbančič, ed., Mojca Bernik, ed., Dunja Mladenčič, ed., Marko Grobelnik, ed., Marjan Heričko, ed., Urban Kordeš, ed., Olga Markič, ed., Ljubljana, Institut "Jožef Stefan", 2008, pp. 174-181.
10. Tomaž Erjavec, "How to encode and display non-Unicode characters in TEI - the case of digital Monumenta frisingensia", In: *Slovo: towards a digital library of South slavic manuscripts: proceedings of the International Conference, 21-26 February 2008, Sofia, Bulgaria*, Heinz Miklas, ed., Anisava Miltenova, Sofia, Bulgarian Academy of Science, Institute of literature, 2008, pp. 194-199.
11. Tomaž Erjavec, Simon Krek, "The JOS morphosyntactically tagged corpus of Slovene", In: *LREC 2008: proceedings*, 6th International Conference on Language Resources and Evaluation, Marrakech, Morocco, May 26 - June 1, 2008, Marrakech, ELRA, 2008, 5 pp.
12. Tomaž Erjavec, Simon Krek, "Oblikoskladenjske specifikacije in označeni korpusi JOS", In: *Zbornik Šeste konference Jezikovne tehnologije, 16. do 17. oktober 2008, Ljubljana, Slovenia: zbornik 11. mednarodne multikonference Informacijska družba - IS 2008, zvezek C: proceedings of the 11th International Multiconference Information Society - IS 2008, volume C*, (Informacijska družba), Tomaž Erjavec, ed., Jerneja Žganec Gros, ed., Ljubljana, Institut "Jožef Stefan", 2008, pp. 49-53.
13. Darja Fišer, Tomaž Erjavec, "Predstavitev in analiza slovenskega wordneta", In: *Zbornik Šeste konference Jezikovne tehnologije, 16. do 17. oktober 2008, Ljubljana, Slovenia: zbornik 11. mednarodne multikonference Informacijska družba - IS 2008, zvezek C: proceedings of the 11th International Multiconference Information Society - IS 2008, volume C*, (Informacijska družba), Tomaž Erjavec, ed., Jerneja Žganec Gros, ed., Ljubljana, Institut "Jožef Stefan", 2008, pp. 37-42.
14. Blaž Fortuna, Dunja Mladenčič, "Using string kernels for classification of Slovenian web documents", In: *From data and information analysis to knowledge engineering: proceedings of the 29th annual conference of the Gesellschaft für Klassifikation e.V., University of Magdeburg, March 9-11, 2005*, (Studies in classification, data analysis, and knowledge organization), Berlin, Springer, 2006, pp. 358-365.
15. Vojko Gorjanc, Simon Krek, Špela Vintar, "Slovene terminology web portal", In: *Proceedings of the XIII EURALEX international congress (Barcelona, 15-19 July 2008)*, [Barcelona, Documenta universitaria, 2008], pp. 971-978.
16. Bryan Griffiths, P. Hallett, Marko Bohanec, G. Osler, B. C. Ball, "A qualitative multi-attribute model for assessing the impact of management options on soil sustainability", In: *Land management in a changing environment: Agriculture and the environment VII: proceedings of the SAC and SEPA biennial conference, Edinburgh, 26-27 March 2008*, Karen Crighton, ed., Rebecca Audsley, ed., [S. l., s. n.], 2008, pp. 63-67.
17. Kristina Hmeljak Sangawa, Tomaž Erjavec, "Črpanje primerov za japonsko-slovenski slovar iz vzporednega korpusa", In: *Zbornik Šeste konference Jezikovne tehnologije, 16. do 17. oktober 2008, Ljubljana, Slovenia: zbornik 11. mednarodne multikonference Informacijska družba - IS 2008, zvezek C: proceedings of the 11th International Multiconference Information Society - IS 2008, volume C*, (Informacijska družba), Tomaž Erjavec, ed., Jerneja Žganec Gros, ed., Ljubljana, Institut "Jožef Stefan", 2008, pp. 33-36.
18. Kristina Hmeljak Sangawa, Tomaž Erjavec, "Gakushoushayou nihongo jisho no tame no taizaku reibun kakutoko", In: *Proceedings of the Workshop on Natural Language Processing for Education, The Fourteenth Annual Meeting of The Association for Natural Language Processing, 21 March 2008, University of Tokyo, [S. l.], The Association for Natural Language Processing, 2008*, pp. 19-22.
19. Matjaž Juršič, Nada Lavrač, "Fuzzy clustering of documents", In: *Zbornik 11. mednarodne multikonference Informacijska družba - IS 2008, 13.-17. oktober 2008: zvezek A: volume A*, (Informacijska družba), Marko Bohanec, ed., Matjaž Gams, ed., Vladislav Rajkovič, ed., Tanja Urbančič, ed., Mojca Bernik, ed., Dunja Mladenčič, ed., Marko Grobelnik, ed., Marjan Heričko, ed., Urban Kordeš, ed., Olga Markič, ed., Ljubljana, Institut "Jožef Stefan", 2008, pp. 178-181.
20. Simon Krek, Vojko Gorjanc, Špela Arhar, "Slovene Terminology Web Portal and the TBX-Compatible Simplified DTD/schema", In: *LREC 2008: proceedings*, 6th International Conference on Language Resources and Evaluation, Marrakech, Morocco, May 26 - June 1, 2008, Marrakech, ELRA, 2008, 5 pp.
21. Alexandra Moraru, Blaž Fortuna, Carolina Fortuna, "Semantic modeling, translation and matching of QoS", In: *Zbornik 11. mednarodne multikonference Informacijska družba - IS 2008, 13.-17.*

- oktober 2008: zvezek A: volume A, (Informacijska družba), Marko Bohanec, ed., Matjaž Gams, ed., Vladislav Rajkovič, ed., Tanja Urbančič, ed., Mojca Bernik, ed., Dunja Mladenec, ed., Marko Grobelnik, ed., Marjan Heričko, ed., Urban Kordeš, ed., Olga Markič, ed., Ljubljana, Institut "Jožef Stefan", 2008, pp. 182-185.
22. Inna Novalija, Dunja Mladenec, "Extending ontologies for annotating business news", In: *Zbornik 11. mednarodne multikonference Informacijska družba - IS 2008, 13.-17. oktober 2008: zvezek A: volume A*, (Informacijska družba), Marko Bohanec, ed., Matjaž Gams, ed., Vladislav Rajkovič, ed., Tanja Urbančič, ed., Mojca Bernik, ed., Dunja Mladenec, ed., Marko Grobelnik, ed., Marjan Heričko, ed., Urban Kordeš, ed., Olga Markič, ed., Ljubljana, Institut "Jožef Stefan", 2008, pp. 186-189.
 23. Dušan Omerčević, Metka Zupančič, Marko Bohanec, Tomaž Kastelic, "Intelligent response to highway traffic situations and road incidents", In: *Greener, safer and smarter road transport for Europe, Proceedings, TRA, Transport Research Arena Europe 2008*, Ljubljana, Slovenia, 21-24 April 2008, Aleš Znidarič, ed., Ljubljana, DDC svetovanje inženiring, ZAG, Zavod za gradbeništvo Slovenije, DRC, Družba v cestni in prometni stroki Slovenije, 2008, pp. 1-6.
 24. Panče Panov, Sašo Džeroski, Larisa N. Soldatova, "OntoDM: an ontology of data mining", In: *ICDM workshop 2008: proceedings*, IEEE International Conference on Data Mining Workshops, 15-19 December 2008, Pisa, Italy, Francesco Bonchi, ed., Danvers, IEEE, 2008, pp. 752-760.
 25. B. Risteovski, S. Loshkovska, Sašo Džeroski, Ivica Slavkov, "A comparison of validation indices for evaluation of clustering results of DNA microarray data", In: *The 2nd International Conference on Bioinformatics and Biomedical Engineering: (ICBBE 2008): May 16-18, 2008, Shanghai, China*, Piscataway, IEEE Xplore, cop. 2008, pp. 587-591.
 26. Jan Rupnik, "Stochastic subgradient approach for solving linear support vector machines: an overview", In: *Zbornik 11. mednarodne multikonference Informacijska družba - IS 2008, 13.-17. oktober 2008: zvezek A: volume A*, (Informacijska družba), Marko Bohanec, ed., Matjaž Gams, ed., Vladislav Rajkovič, ed., Tanja Urbančič, ed., Mojca Bernik, ed., Dunja Mladenec, ed., Marko Grobelnik, ed., Marjan Heričko, ed., Urban Kordeš, ed., Olga Markič, ed., Ljubljana, Institut "Jožef Stefan", 2008, pp. 194-197.
 27. Jan Rupnik, Miha Grčar, Tomaž Erjavec, "Improving morphosyntactic tagging of Slovene by tagger combination", In: *Zbornik Šeste konference Jezikovne tehnologije, 16. do 17. oktober 2008, Ljubljana, Slovenia: zbornik 11. mednarodne multikonference Informacijska družba - IS 2008, zvezek C: proceedings of the 11th International Multiconference Information Society - IS 2008, volume C*, (Informacijska družba), Tomaž Erjavec, ed., Jerneja Žganec Gros, ed., Ljubljana, Institut "Jožef Stefan", 2008, pp. 110-115.
 28. Delia Rusu, Blaž Fortuna, Marko Grobelnik, Dunja Mladenec, "Semantic graphs derived from triplets with application in document summarization", In: *Zbornik 11. mednarodne multikonference Informacijska družba - IS 2008, 13.-17. oktober 2008: zvezek A: volume A*, (Informacijska družba), Marko Bohanec, ed., Matjaž Gams, ed., Vladislav Rajkovič, ed., Tanja Urbančič, ed., Mojca Bernik, ed., Dunja Mladenec, ed., Marko Grobelnik, ed., Marjan Heričko, ed., Urban Kordeš, ed., Olga Markič, ed., Ljubljana, Institut "Jožef Stefan", 2008, pp. 198-201.
 29. Serge Sharoff, Mikhail Kopotev, Tomaž Erjavec, Anna Feldmann, Dagmar Divjak, "Designing and evaluating a Russian tagset", In: *LREC 2008: proceedings*, 6th International Conference on Language Resources and Evaluation, Marrakech, Morocco, May 26 - June 1, 2008, Marrakech, ELRA, 2008, 7 pp.
 30. Alenka Stanič Lang, Tanja Urbančič, "Izobraževanje danes za poklice prihodnosti: izzivi informacijske družbe: challenges of information society", In: *Vzgoja in izobraževanje v informacijski družbi: zbornik referatov: conference proceedings*, Vladislav Rajkovič, ed., Tanja Urbančič, ed., Mojca Bernik, ed., Kranj, Fakulteta za organizacijske vede, 2008, pp. 298-304.
 31. Alenka Stanič Lang, Tanja Urbančič, "Poklici prihodnosti v luči trajnostnega razvoja", In: *Znanje za trajnostni razvoj: zbornik 27. mednarodne znanstvene konference o razvoju organizacijskih znanosti, Slovenija, Portorož, 19.-21. marec 2008: proceedings of the 27th International Conference on Organizational Science Development, Slovenia, Portorož, March, 19th-21th, 2008*, 27. mednarodna znanstvena konferenca o razvoju organizacijskih znanosti, Slovenija, Portorož, 19.-21. marec 2008, Vladislav Rajkovič, ed., Eva Jereb, ed., Tomaž Kern, ed., Miroljub Kljajić, ed., Milan Pagon, ed., Goran Vuković, ed., Kranj, Moderna organizacija, 2008, pp. 2547-2552.
 32. Dan Tufiş, Svetla Koeva, Tomaž Erjavec, Maria Gavrilidou, Cvetana Krstev, "Building language resources and translation models for machine translation focused on South nad Balkan Languages", In: *Proceedings, The Sixth International Conference Formal Approaches to South Slavic and Balkan Languages*, 25-28 September 2008, Dubrovnik, Croatia, Zagreb, Croatian Language Technologies Society, Faculty of Humanities and Social Sciences, 2008, pp. 145-152.
 33. Špela Vintar, Tomaž Erjavec, "iKorpus in luščenje izrazja za Islovar", In: *Zbornik Šeste konference Jezikovne tehnologije, 16. do 17. oktober 2008, Ljubljana, Slovenia: zbornik 11. mednarodne multikonference Informacijska družba - IS 2008, zvezek C: proceedings of the 11th International Multiconference Information Society - IS 2008, volume C*, (Informacijska družba), Tomaž Erjavec, ed., Jerneja Žganec Gros, ed., Ljubljana, Institut "Jožef Stefan", 2008, pp. 65-69.
 34. Monika Žáková, Petr Křemen, Filip Železný, Nada Lavrač, "Using ontological reasoning and planning for data mining workflow composition", In: *Third generation data mining: towards service-oriented knowledge discovery*, European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases, EMM/PPKDD 2008, 15-19 September, 2008, Antwerp, Belgium, Nada Lavrač, ed., Joost N. Kok, ed., Jeroen de Bruin, ed., Vid Podpečan, ed., [S. l., s. n.], 2008, pp. 35-41.

TEXTBOOKS AND LECTURE NOTES

1. Nataša Atanasova, Marko Debeljak, Sašo Džeroski, Andrej Kobler, Boris Kompare, Ljupčo Todorovski, Bernard Ženko, *Analysis of environmental data with machine learning methods: Ljubljana, March 17-21, 2008. Part 2*, Ljubljana, Jožef Stefan Institute, Center for Knowledge Transfer in Information Technology, 2008.
2. Nataša Atanasova, Marko Debeljak, Sašo Džeroski, Andrej Kobler, Boris Kompare, Ljupčo Todorovski, Bernard Ženko, *Analysis of environmental data with machine learning methods: Ljubljana, March 17-21, 2008. Part 3*, Ljubljana, Jožef Stefan Institute, Center for Knowledge Transfer in Information Technology, 2008, Systemske zahteve: Adobe Acrobat Reader.
3. Ivan Bratko, Sašo Džeroski, *Analysis of environmental data with machine learning methods: Ljubljana, March 17-21, 2008. Part 1*, Ljubljana, Jožef Stefan Institute, Center for Knowledge Transfer in Information Technology, 2008.

THESES

B. Sc. Theses

1. Blaž Novak, *Odkrivanje tematik v zaporedju besedil in sledenje njihovim spremembam: diplomatska naloga*, Ljubljana, [B. Novak], 2008.
2. Mitja Trampuš, *Miselni poker: diplomatska naloga*, Ljubljana, [M. Trampuš], 2008.