

Deveta konferenca JEZIKOVNE TEHNOLOGIJE
9. – 10. oktobra 2014

Ninth LANGUAGE TECHNOLOGIES Conference
9th – 10th October 2014

ČETRTEK, 9. oktober
THURSDAY, October 9th

Predavalnica: Predavalnica Mednarodne podiplomske šole Jožefa Stefana
Venue: Lecture hall of the Jožef Stefan International Postgraduate School

8:50 – 10:30

- Otvoritev / Opening Address

Predsednik / Chair: Tomaž Erjavec

- Vabljeni predavanja / Invited Lecture
Maciej Piasecki, Tehnična univerza v Wroclawu
User-driven Language Technology Infrastructure – the Case of CLARIN-PL
- Vabljeni predavanja / Invited Lecture
Matej Āurčo, Avstrijska akademija znanosti
CLARIN-DARIAH.AT – Weaving the network

10:30 - 11:00 Odmor za kavo / Coffee break

11:00 – 12:20

Predsednik / Chair: Špela Vintar

- Tomaž Erjavec, Jan Jona Javoršek, Simon Krek
Raziskovalna infrastruktura CLARIN.SI / The research infrastructure CLARIN.SI
- Nikola Ljubešić, Kaja Dobrovoljc, Simon Krek, Marina Peršurić, Darja Fišer
hrMWElex – a MWE lexicon of Croatian extracted from a parsed gigacorpus
- Petra Almić, Jan Šnajder
Determining the Semantic Compositionality of Croatian Multi-Word Expressions
- Cvetana Krstev, Staša Vujić Stanković, Duško Vitas
Approximate Measures in the Culinary Domain: Ontology and Lexical Resources

12:20 - 13:40 Odmor za kosilo / Lunch break

13:40 – 15:00

Predsednik / Chair: Jan Šnajder

- Darja Fišer, Benoît Sagot
Avtomatska razširitev in čiščenje sloWNeta
- Tomaž Erjavec, Nikola Ljubešić
The slWaC 2.0 Corpus of the Slovene Web
- Darja Fišer, Tomaž Erjavec, Ana Zwitter Vitez, Nikola Ljubešić
Janes se predstavi: metode, orodja in viri za nestandardno slovenščino
- Filip Klubička, Nikola Ljubešić
Using crowdsourcing in building a morphosyntactically annotated and lemmatized silver standard corpus of Croatian

15:00 - 15:30 Odmor za kavo / Coffee break

15:30 – 17:50

Predsednik: Darja Fišer

- Leo Zuanović, Mladen Karan, Jan Šnajder
Experiments with Neural Word Embeddings for Croatian
- Jelena Jaćimović, Cvetana Krstev, Drago Jelovac
Automatic de-identification of protected health information
- Kaja Dobrovoljc
Procesiranje slovenskega jezika v razvojnem okolju NooJ
- Krešimir Baksa, Dino Dolović, Goran Glavaš, Jan Šnajder
Named Entity Recognition in Croatian Tweets
- Nikola Ljubešić, Denis Kranjčić
Discriminating between VERY similar languages among Twitter users
- Siniša Biđin, Jan Šnajder, Goran Glavaš
Predicting Croatian Phrase Sentiment Using a Deep Matrix-Vector Model
- Luka Skukan, Goran Glavaš, Jan Šnajder
HeidelTime.Hr: Extracting and Normalizing Temporal Expressions in Croatian

Deveta konferenca JEZIKOVNE TEHNOLOGIJE
9. – 10. oktobra 2014

Ninth LANGUAGE TECHNOLOGIES Conference
9th – 10th October 2014

PETEK, 10. oktober
FRIDAY, October 10th

Predavalnica: Predavalnica Mednarodne podiplomske šole Jožefa Stefana
Venue: Lecture hall of the Jožef Stefan International Postgraduate School

09:00 – 10:30

Predsednik / Chair: Jerneja Žganec Gros

- Vabljeni predavanja / Invited Lecture
Maja Popović, DFKI
Machine translation: approaches, applications and challenges
- Kristijan Armeni, Grega Repovš, Špela Vintar
Eye tracking measures as indicators of readability for machine-translated texts
- Mirjam Sepesy Maučec
Evalvacija slovensko-srbskih strojnih prevodov v projektu SUMAT

10:30 - 11:00 Odmor za kavo / Coffee break

11:00 – 12:40

Predsednik / Chair: Simon Krek

- Senja Pollak, Biljana Božinovski
Luščenje borzne terminologije
- Mario Jurišić, Špela Vintar
Analiza uporabe slovnčnih pregledovalnikov za slovenščino
- Miro Romih
SecondEGO – virtualni pomočnik za vsakogar
- Ana Zwitter Vitez
Ugotavljanje avtorstva besedil: primer »Trenirkarjev«
- Peter Holozan
Razreševanje sklicev pri analizi slovenskih besedil

12:40 - 13:40 Odmor za kosilo / Lunch break

13:40 – 15:00

Predsednik / Chair: France Mihelič

- Boštjan Vesnicher, Jerneja Žganec Gros, Simon Dobrišek, Vitomir Štruc
Alp-ULj Speaker Recognition System for the NIST 2014 i-Vector Challenge
- Andrej Žgank, Gregor Donaj, Mirjam Sepesy Maučec
Razpoznavnik tekočega govora UMB Broadcast News 2014: kakšno vlogo igra velikost učnih virov?
- Darinka Verdonik
Vprašanja zapisovanja govora v govornem korpusu Gos
- Martina Ozbič, Damjana Kogovšek, Mojca Muznik
Variabilnost izgovora kot ovira pri avtomatskem prepoznavanju govora: primer epenteze, epiteze in proteze v govoru slovenskih predšolskih otrok

15:00 - 15:30 Odmor za kavo / Coffee break

15:30 – 16:30

Predsednik / Chair: Darinka Verdonik

- Špela Vintar, Boštjan Jerko, Marjetka Kulovec
Prvi leksikalni podatki o slovenskem znakovnem jeziku iz korpusa Signor
- Tadej Justin, France Mihelič and Janez Žibert
Razvoj zbirke slovenskega emocionalnega govora iz radijskih iger - EmoLUKS
- Žiga Golob, Jerneja Žganec Gros, Simon Dobrišek
Končni super pretvorniki za predstavitev slovarjev izgovorjav pri sintezi govora

17:00 – 18:30

Skupščina SDJT

General Assembly of the Slovenian Language Technologies Society