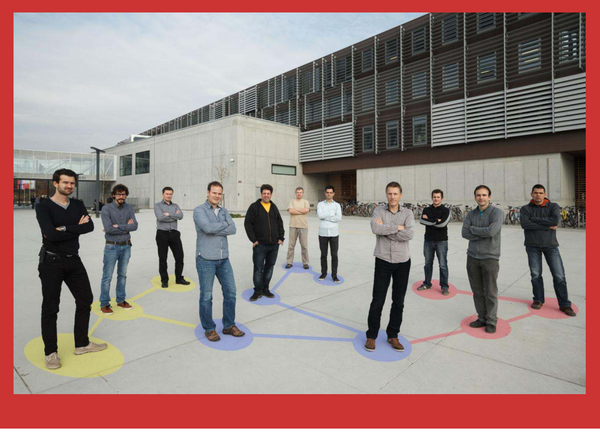
Univerza *v Ljubljani*
Fakulteta *za računalništvo in informatiko*

LPT

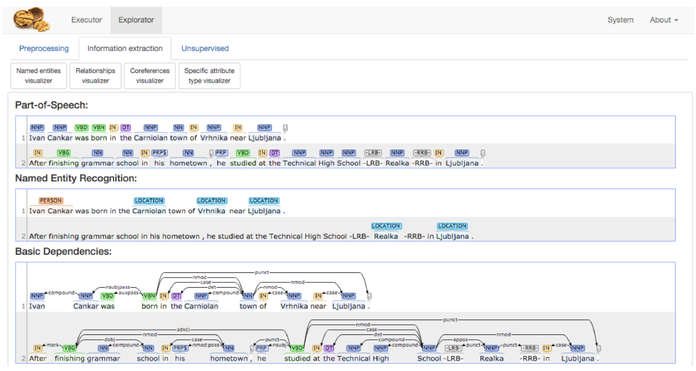
Laboratorij za podatkovne tehnologije
Laboratory for data technologies

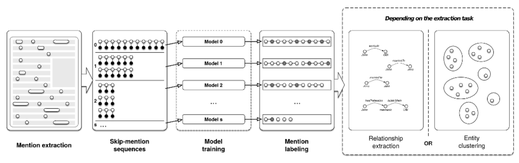


Information extraction

Žitnik, et al. (2015). Sieve-Based Relation Extraction of Gene Regulatory Networks from Biological Literature, BMC Bioinformatics
Žitnik, Bajec (2015). Iterative joint extraction of entities, relationships and coreferences from text sources, RCIS 2015
Žitnik, Šubelj, Bajec (2014). SkipCor: Skip-Mention Coreference Resolution using Linear-Chain Conditional Random Fields, PLoS One
Žitnik, Bajec M. (2013). Text Mining in Medicine, Book chapter

NutIE - *web-based information extraction*





Slovene language processing:
Open information extraction
Coreference resolution

Conversational corpora mining

Named entity recognition
Relationship extraction

Discourse sense classification from scratch

CONLL 2016 Shared Task on English/Chinese

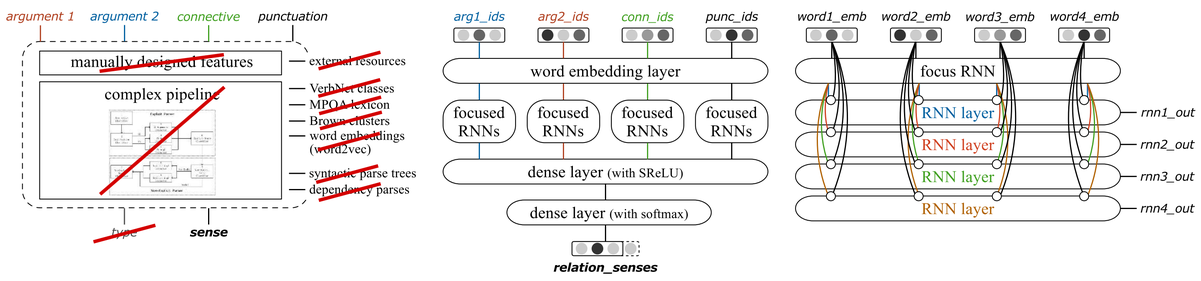
- locate explicit or implicit discourse **connective**
- locate text spans for **argument 1** and **argument 2**
- predict discourse relation **sense** (4 classes, 16 types, 23 subtypes)

A = Implicit : Comparison.Contrast
B = Explicit : Comparison.Concession
C = Explicit : Contingency.Condition

According to Lawrence, "Kemper is using program trading." He added that "having just one such firm doesn't matter. But if there are more, then it may be important."

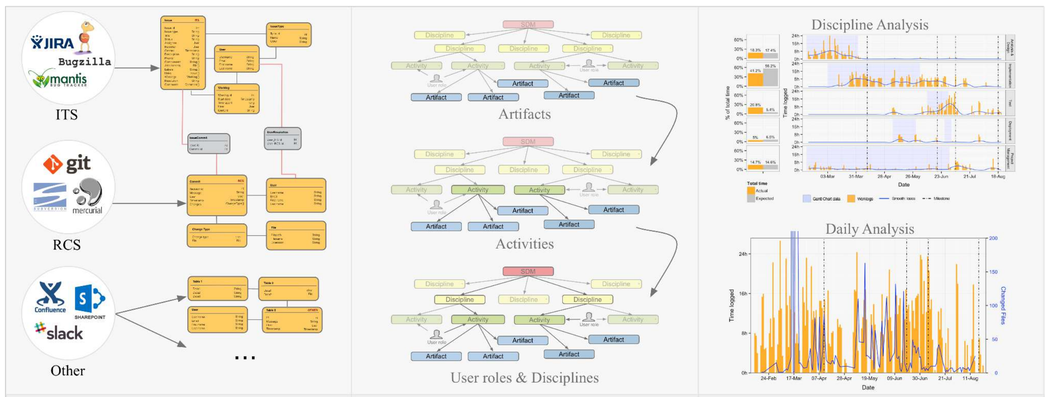
A [arg1] ..... [arg2]
B ..... [arg1] [ ][arg2]
C ..... [ ][ ][arg1] [ ][arg2]

From scratch with **deep neural networks** (focused RNNs)



Weiss, Bajec (2016), in Proc. of the **CONLL 2016**, pp. 50–54.

Mining Software Repositories



Prepare and link data from software repositories

Reconstruct software development method/process

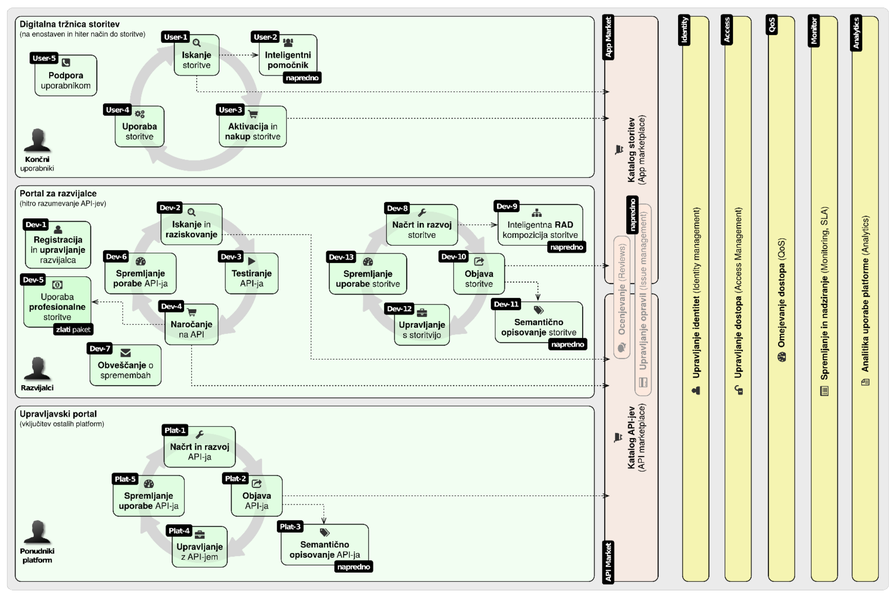
Project/process performance analysis

Cooperation

- Semi-automatic improvement of software development methods: doctoral consortium paper (RCIS 2013) - **best paper award**
- Intelligent agile method framework (ENASE 2013)
- Comparison of software repositories for their usability in software process reconstruction (RCIS 2015)
- Software process reconstruction toolset: http://ispr.jmlabs.eu

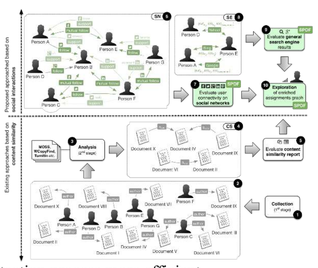
Publications

Integracijska platforma **EkoSmart**



Plagiarism detection with social data

A novel approach focussing on social aspects of potential plagiarists, by taking into account their social network connections, activities, and information from the Web, to support investigator's work in the third and the fourth stage of FSPDP, thereby making the plagiarism detection process more efficient.

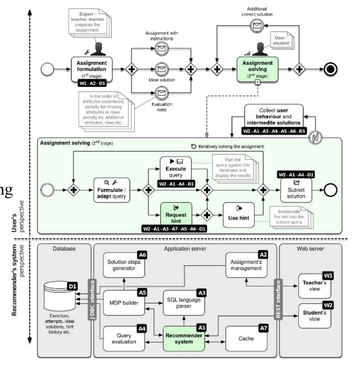


ZRNEC, Aljaž, LAVBIČ, Dejan. ***Social network aided plagiarism detection****. British journal of educational technology, 48(1), 2017, pp. 113-128.*

ZRNEC, Aljaž, LAVBIČ, Dejan. ***The role of social connections in plagiarism detection****. Learning technology for education in cloud: LTEC 2015, Maribor, Slovenia, August 24-28, 2015, Springer proceedings, pp. 54-63.*

Intelligent SQL tutor

The adaptive design of the proposed approach aids users in learning SQL by supporting their own path to the solution and employing successful previous attempts, while not enforcing the ideal solution provided by the instructor.

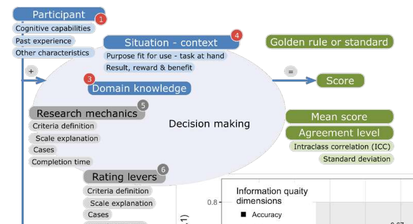


LAVBIČ, Dejan, MATEK, Tadej, ZRNEC, Aljaž. ***Recommender system for learning SQL using hints****. Interactive Learning Environments, doi:10.1080 / 10494820.2016.1244084, 2016.*

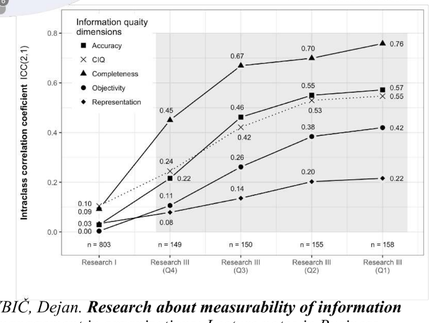
MATEK, Tadej, ZRNEC, Aljaž, LAVBIČ, Dejan. ***Learning SQL with artificial intelligent aided approach****. ICETC 2016, September 28-30, 2016, best paper award.*

Improving information quality

Identify and investigate key drivers of information quality improvement.



Increased agreement level and improved information quality, highly correlated with "task at hand".



http://bit.ly/InfQ-game
Participate in research

FIDLER, Miloš, LAVBIČ, Dejan. ***Research about measurability of information quality****. Knowledge management in organizations: Lecture notes in Business Information processing, Springer, 2015, pp. 272-281.*

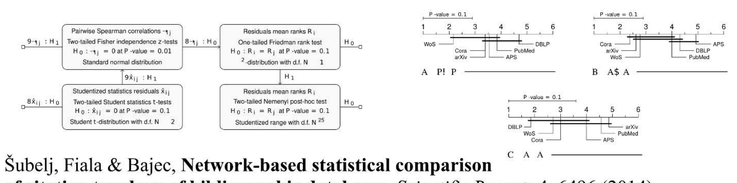
FIDLER, Miloš, LAVBIČ, Dejan. ***Improving Information Quality of online articles with cooperative principle****. Online Information Review, 2016, under review.*

Dynamic Business Process Modelling

VASILECAS, Olegas, KALIBATIENE, Diana, LAVBIČ, Dejan. ***Rule- and Context-Based Dynamic Business Process Modelling and Simulation****. Journal of Systems and Software, 122, 2016.*

Reliability of Bibliographic Databases

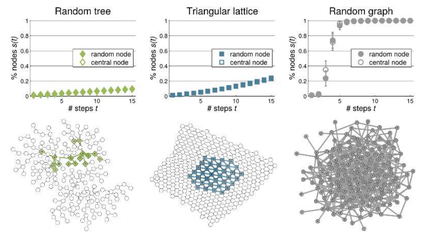
Are WoS and SCOPUS more reliable than DBLP?



Šubelj, Fiala & Bajec, **Network-based statistical comparison of citation topology of bibliographic databases**, *Scientific Reports* 4, 6496 (2014)
Šubelj, Bajec, Boshkoska, Kastrin & Levnajić, **Quantifying the consistency of scientific databases**, *PLoS ONE* 10(5), e0127390 (2015)

Convexity in Networks

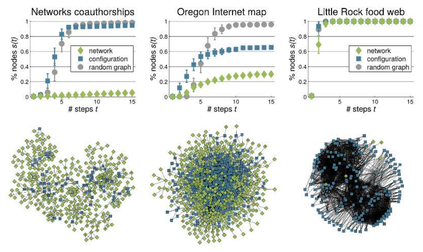
Are real networks convex or not?



Clustering of Scientific Publications

Which clustering method reveals research topics of publications?

class	method	description
Spectral analysis	Graclus(S|L)	*k*-means clustering ↑iteration
	METIS(S|L)	multi-level *k*-way partitioning
Map equation	Infomap	information flows compression
	Hiermap	hierarchical flows compression
Modularity optimization	Louvain	greedy hierarchical optimization
	Mouvain	multi-level hierarchical optimization
	SLM	smart local moving optimization
Statistical methods	OSLOM	order statistics local optimization method
Label propagation	LPA	label propagation algorithm
	BPA	balanced propagation algorithm
	DPA	diffusion-propagation algorithm
	HPA	hierarchical propagation algorithm
	COPRA	community overlap propagation algorithm
Random walks	Walktrap	random walks hierarchical clustering
Link clustering	Links(S|L)	link similarity hierarchical clustering
Graph models	BigClam(S|L)	cluster affiliation matrix factorization
	CoDA(S|L)	communities through directed affiliations
Ego-networks	DEMON	democratic estimate of modular organization
Cliques	SCP	sequential clique percolation
	GCE	greedy clique expansion
2-step methods	Metilus	METIS+Graclus
	Gracmap	Graclus+Infomap
	METIS+Infomap	METIS+Infomap
	Louvmap	Louvain+Infomap
	Labmap	LPA+Infomap



Marc & Šubelj, **Convexity in complex networks**, submitted to *Network Science*, pp. 26 (2016)
Šubelj & Marc, **On convexity in complex networks**, In: *Proceedings of CompleNet* (Dubrovnik, 2017)

Šubelj, Van Eck & Waltman, **Clustering scientific publications based on citation relations**, *PLoS ONE* 11(4), e0154404 (2016)
Šubelj, Van Eck & Waltman, **Comparison of methods for clustering citation networks**, In: *Proceedings of NetSci-X* (Wroclaw, 2016)

Software Workbench for Interactive, Time Critical and Highly self-adaptive Cloud applications

SWITCH

Project objectives

Improve development and execution model of time critical applications in the programmable, elastic and visualised environment.

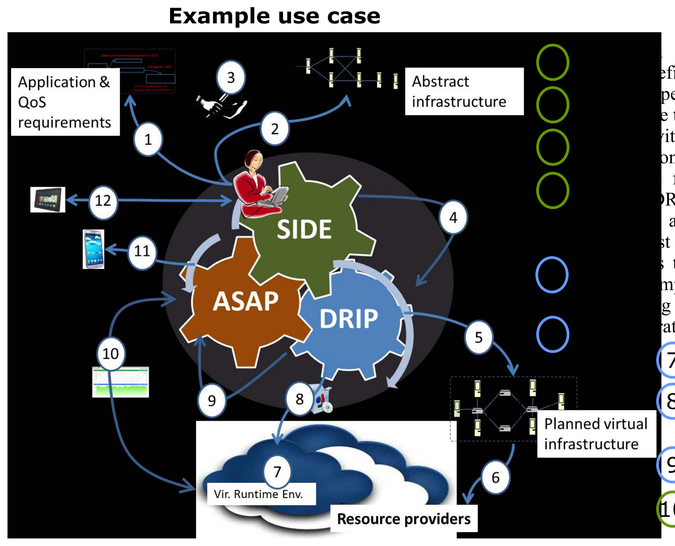
Increase productivity and decrease operational costs of application development for time critical applications.

Improve deployment efficiency for time critical applications in Clouds by affecting Cloud providers with supplying and inducing their services to the customers.

Decrease operational costs of time critical services.

Promote competitiveness of companies by providing time-critical mechanisms for negotiation and time-critical software services to optimize the performance of Cloud services.

Example use case



developer composes the application
efines the QoS constraints.
per specifies an abstract network over-
e the runtime environment.
ities are optimised and aided using a
oning component.
from step 1 and 2 are passed from
RIP; the developer specifies require-
as resource providers to be used, and
t budget for application execution.
the concrete virtual runtime environ-
mputing, storage and network elements
g with the app-level QoS constraints.
ates SLA with the resource providers.

7 The resource providers provision the virtual environment.
8 DRIP customises the virtual environment and deploys necessary services for the application.
9 DRIP unbundles and executes the application.
10 At runtime, SIDE allows the user to query and visualise the runtime status of the application and runtime environment.
11 At runtime, receive notification of system status and inform the user.
12 Directly manipulate the system execution and expose real-time monitoring information from the database, generated by ASAP.

Project partners

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 643963 (SWITCH project).